The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Burke, Terence (1984) The ecological genetics of two populations of the house sparrow, Passer domesticus. PhD thesis, University of Nottingham.

Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/14501/1/350190.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

- Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners.
- To the extent reasonable and practicable the material made available in Nottingham ePrints has been checked for eligibility before being made available.
- Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
- Quotations or similar reproductions must be sufficiently acknowledged.

Please see our full end user licence at:
http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

THE ECOLOGICAL GENETICS OF TWO POPULATIONS OF THE HOUSE SPARROW, *PASSER DOMESTICUS*.

Terence Burke B.Sc.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy, July 1984

BRITISH LIBRARY HSIRB

IMAGING SERVICES NORTH

Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
www.bl.uk

VOLUME CONTAINS CLEAR OVERLAYS

OVERLAYS HAVE BEEN SCANNED SEPERATELY AND THEN AGAIN OVER THE RELEVANT PAGE

CONTENTS

ABSTRACT

The biochemical genetics of two natural populations of house sparrows, Passer domesticus, at sites 20km apart in Nottinghamshire, England, were investigated. Seven polymorphic protein loci were sampled non-destructively by taking blood samples from over 1500 individually marked birds. A detailed investigation of the genetics of these loci was conducted for 124 clutches containing 357 nestlings where the parents were also sampled. Segregations at four loci (6PGD, PEPD2, PEPD3 and IDHC) agreed with a simple Mendelian model of codominant inheritance. One locus (EST2) contained null alleles. Two loci (PEPD3 and GP1) showed segregation distortion in all sex, site and year classes. This distortion was not attributable to the misinterpretation of gel patterns; possible causes involving the operation of natural selection were discussed. Linkage analyses were conducted, and no significant evidence was obtained for linkage between any combination of loci.

Of the nestling genotypes, 12.9% were interpreted as being genetically incompatible with those of their parents. Exclusion probabilities were calculated as 43-51% for nonpaternity and 59-67% for nonparentage. The applicability of these estimated probabilities was tested by the random reassortment and comparison of observed parental genotypes among observed sibship genotypes. Significantly fewer nestlings were excluded in these simulations than expected from calculated exclusion probabilities, though the distribution of multiple mismatches did not differ from expectation. A deficiency of multiple mismatches was found in the field data, implying the occurrence of errors; the possible sources of error were considered. The most parsimonious interpretation of those mismatches that did not appear to be due to errors was that they resulted from a rate of nonpaternity of about 6%. No heterogeneity in the rate of mismatches was observed within or among breeding seasons or sites.

Genotype and allele frequencies were presented for each

locus in each age, sex and sampling year class at each study site. The samples were not found to depart from Hardy Weinberg equilibrium, and there was no evidence for significant inbreeding within sites. There were no differences in allelic distributions between the sexes or among years for adults within the populations. No differences were found among age groups or nestling year classes when allowance was made for sib correlations. Heterozygosities were higher at Brackenhurst than at Sutton Bonington for most loci, and the overall difference was significant. There was a particularly large difference in allele frequencies between nestlings in each population for GPl. Digenic gametic disequilibria were investigated.

A detailed analysis of the mating types was made. No evidence was obtained for any departure from random mating at the protein loci. There was a significant tendency amongst the loci and samples for the inbreeding coefficients of the successful breeders to be negative. Significant assortative mating was found with respect to weight and tail-length in one population.

ACKNOWLEDGEMENTS

The study was initiated at the suggestion of Dr. David Parkin, whose interest, guidance and encouragement I have enjoyed throughout. The work was undertaken in the Genetics Department at Nottingham and I thank Professor Bryan Clarke for the provision of excellent facilities.

The people living and working at Brackenhurst College and the University of Nottingham's Experimental Farm at Sutton Bonington are thanked for their friendly interest and enthusiastic cooperation.

Members of the Genetics Department contributed to the study in a variety of ways. I especially thank Steve Cole for teaching me about gelling, and Jon Lewis for technical assistance way beyond the call of duty.

Advice on statistical computing packages was provided by Mr. P. Riley. Dr. Peter G.H. Evans showed me a method for taking blood samples, and provided one of his nestbox traps.

The assistance of members of the South Notts. Ringing Group was freely given on numerous occasions, and is gratefully acknowledged: Mick Thompson and Dave Pons are particularly thanked.

Many other friends helped at crucial times: Allyn Thomas occasionally helped with fieldwork in 1980; Ian Masters discovered the joys of mist-netting at sunrise, and provided invaluable field assistance in 1981; Peter Shelton, John Stephen, and especially Des and Dawn Thompson and my wife Sarah provided encouragement, both spiritual and material, in finally getting me to the binder's.

This project was supported by the Natural Environment Research Council.

CHAPTER 1

INTRODUCTION

1.1 General Introduction

Ecological genetics is the study of genetic processes in natural populations of organisms. The quest is for an understanding of the forces moulding the genetic constitution of natural populations. Such an understanding is of obvious importance to the study of the process of evolution. The focus of ecological genetics is, however, on evolutionary phenomena and not necessarily on evolutionary progression (Ford 1975).

A major feature of the genetic makeup of interbreeding populations is the genetic variability found within those populations. Ecological genetic studies have been aimed particularly at understanding this variability. All genetic variability is ultimately encoded as differences in the DNA of individuals. Whilst it is now feasible to identify genetic differences among organisms within a population by the direct characterisation of their DNA (e.g. Kan and Dozy 1980), almost all ecological genetic studies have been concerned with phenotypic variation that has a genetic component.

Electrophoretically detected variants of soluble proteins are one class of phenotypic variation, and are generally associated with differences at individual genomic loci. This thesis is particularly concerned with such variation in populations of the house sparrow, Passer domesticus. Some consideration is also made of metrical variation, believed to be under polygenic control. The genetic variability is itself used in investigations regarding the mating system. A detailed knowledge of mating and population structure is

fundamental to an understanding of genetic processes in natural populations.

1.2 Studies of Protein Polymorphisms

Gel electrophoresis allows the separation of soluble protein molecules that differ in size, shape or charge. Such proteins migrate through a gel at different rates when an electric current is applied; the rate of migration depends upon the voltage applied across the gel, the concentration of the gel substrate, the pH and the ionic concentration of the gel buffer. At the end of the electrophoretic run the proteins are visualised as bands, either by direct staining or, in the case of enzymes, through staining reactions designed to produce a detectable product. Enzyme variants are described as isozymes (or allozymes), and their banding patterns as zymograms.

A mutation at a structural protein locus (such as one coding for an enzyme) may result in the substitution of an amino acid which in turn may result in a change in the shape or charge of the protein. Thus, many variants at soluble protein loci are detectable electrophoretically.

Paper electrophoresis techniques allowed the study of protein variation of egg albumins and haemoglobins as early as 1938 (Landsteiner et al. 1938). It was not until the advent of starch gel electrophoresis in the 1950s that such studies started to become commonplace (e.g. Lush 1961, Mueller et al. 1962). Throughout this period all such variation was regarded by investigators as likely to be adaptive. Consequently much of the early effort was directed towards detecting potential economically important fitness differences among protein phenotypes in domestic animal breeds and

many correlations with a variety of components of fitness were reported (e.g. Morton et al. 1965, Fowle et al. 1967).

Two extreme views concerning the degree of genetic variation in natural populations were current at the time. The 'classical' view as it became known (Dobzhansky 1955), propounded particularly by Muller (1962), argued that genetic variation would be extremely limited. New mutants would either be selectively disadvantageous and be quickly eliminated, never having exceeded a low frequency (purifying selection), or else advantageous, and quickly driven by natural selection to fixation (directional selection). The alternative hypothesis (e.g. Wallace 1958) predicted that almost all loci in an individual would be heterozygous; hence there would be immense variability with many alleles at every locus. This became known as the 'balance' view, following Dobzhansky's proposition that most polymorphisms were maintained by balancing selective forces (Dobzhansky 1955). The argument became less abstract when electrophoretic techniques were applied to the study of a large number of enzyme loci in Drosophila pseudoobscura (Hubby & Lewontin 1966) and in our own species (Harris 1966).

These and the many later studies demonstrated that the estimated level of heterozygosity (the proportion of an individual's loci in the heterozygous condition) for soluble protein loci was typically of the order of 10%. This was an immensely greater proportion than could be accounted for under the classical hypothesis. The heterozygosity estimates would themselves be considerably less than the true level of heterozygosity, as only about 30% of amino acid substitutions are expected to be electrophoretically detectable. Whether the studied loci are truly representative of the genome as a whole remains unknown. Consistent

differences among the different species studied for the heterozygosity of different classes of soluble proteins (Gillespie and Langley 1980, O'Brien et al. 1980) show that different levels of heterozygosity are attainable. Some workers have used O'Farrell's (1975) two-dimensional gelling technique in order to greatly increase the number of loci surveyed, and thereby test the generality of extrapolations made from enzyme studies.

The interpretation of these data remains controversial: studies in man, Drosophila, pigeons, and Mus reported that heterozygosity in the larger samples of loci was significantly less ($\bar{H} \simeq 0.02$ to 0.04) than in enzyme studies (Walton et al. 1979, Leigh Brown and Langley 1979, Schenker 1979, Racine and Langley 1980). McLellan et al. (1983) have, however, argued that 2-D electrophoresis is less sensitive to the detection of substitutions than 1-D electrophoresis, and that this will account, in part at least, for the suggested differences in heterozygosity between enzymes and other proteins.

Whatever the exact level of variability, it is certainly substantially greater than could satisfactorily be explained by the traditional view of balancing selection due principally to heterozygote advantage. Kimura and Crow (1964) showed that the expected mean equilibrium level of heterozygosity ($\bar{H}$) at loci subject to mutation and random genetic drift, in the absence of selection, is given by the relationship:

$$\bar{H} = 4N_e v/(4N_e v + 1)$$

where N_e is the effective population size and v is the mutation rate per locus per generation. This ultimately led to the proposal of neutral theory (see Kimura 1983), asserting that the great majority of evolutionary changes at the molecular level, such as those

producing enzyme variants, were not subject to Darwinian natural selection. Selection would still operate on deleterious or advantageous mutants in the same manner as envisaged by the classical theory. Hence neutralism came to be described as the 'neoclassical' hypothesis (Lewontin 1974). Kimura's suggestions were supported by King and Jukes (1969) who sensationalized the debate by asserting that most evolution at the molecular level was 'non-Darwinian'. As pointed out by Grant (1977), however, Darwin did not assert that all phenotypic variability would necessarily be subject to natural selection (Darwin 1859).

Thus the discovery of large amounts of genetic variation did not resolve the original debate concerning the nature of selective forces acting at the molecular level; in this sense the 'new' debate, the neutralism-selectionism controversy, was a continuation of the old classical-balance one. (See Lewontin 1974 and Crow 1981 for reviews).

The availability of electrophoretic techniques and their ready applicability to population studies led to a heightening of the debate as many more investigators collected field data. The question of the extent to which enzyme polymorphisms are maintained by selective or else neutral stochastic forces remains unresolved. Despite the intensity of investigation, little convincing evidence for the importance of selection has been accummulated. On the other hand, the predictions of the neutral theory are neither disproven nor fully vindicated. For example, from Kimura and Crow's equation (above) relating heterozygosity to effective population size and mutation rate, a correlation is predicted between heterozygosity and effective population size. The relative similarity of heterozygosities in a wide range of organisms whose N_e's must be

vastly different makes such a correlation extremely unlikely.

Several mechanisms through which balancing selection might maintain genetic polymorphisms have been proposed. These include heterozygote advantage, as already referred to (e.g. Lewontin et al. 1978), frequency-dependent selection (e.g. Clarke and O'Donald 1964, Kojima and Yarbrough 1967) forces of directional selection balanced across sexes, age classes, space or time (e.g. Hedrick et al. 1976). If selection is involved in the maintenance of all, or even a large proportion, of polymorphisms, then it is clear that it cannot act constantly and independently on all loci. Franklin and Lewontin (1970) proposed that selection acts on coadapted gene complexes, and predicted that widespread linkage disequilibria (gametic phase disequilibria) would develop amongst alleles at different loci. Almost all the data concerning linkage disequilibria in wild populations have been obtained from studies of Drosophila (e.g. Loukas et al. 1980), and have not supported this prediction. If selection is operating independently on such large numbers of loci, then it must operate in a density-dependent (or threshold) fashion upon many loci simultaneously (e.g. King 1967). Wills (1978, 1980) proposed that all variation could be maintained by a kind of truncation selection termed 'rank-order' selection.

Whilst fitness differences may be observed for biochemically detectable genotypes at a locus, it is more difficult to prove that selection is acting directly on that locus, and not on one in linkage disequilibrium with it (Clarke 1975). The sickle-cell haemoglobin polymorphism (Allison 1964) is perhaps the only well understood example of the direct selective maintenance of a polymorphism. This polymorphism is maintained by heterozygote advantage associated with resistance to malaria. There is strong,

but less conclusive, evidence for the direct selective maintenance of other polymorphisms, such as those others of haemoglobin and G6PD associated with malaria resistance in humans (see Wills 1980), and the warfarin resistance polymorphism in rats, Rattus norvegicus (Bishop, Hartley and Partridge 1977). Heterosis has also been implicated in the maintenance of these polymorphisms.

Clarke (1975) emphasized the importance of knowing the physiological role of allozymes to an understanding of the possible selective maintenance of enzyme polymorphisms. Koehn et al. (1983) have recently reviewed those cases where such an understanding appears to be emerging, but consider that none can as yet be regarded as definitive. They include studies of alcohol dehydrogenase (ADH), α-amylase, esterase-6, α-glycerophosphate dehydrogenase and glucose-6-phosphate dehydrogenase/6-phosphogluconate dehydrogenase polymorphisms in Drosophila, the lactate dehydrogenase polymorphism in the teleost fish Fundulus heteroclitus, a glutamate-pyruvate transaminase polymorphism in the intertidal copepod Tigriopus californicus, and the aminopeptidase I. polymorphism in Mytilus edulis. In all these cases activity differences (due to concentration or catalytic activity) have been found among allozymes, together with strong laboratory evidence of fitness differences and/or field evidence for the action of natural selection upon genotypes. For example, the allozyme produced by the Drosophila melanogaster $ADH^{F/F}$ genotype exhibits greater activity than that of $ADH^{S/S}$ (Day et al. 1974, Zera et al. 1983). ADH allele frequencies in the wild vary clinally in a similar way on different continents, and there is strong evidence that they are maintained by selective environmental gradients (Oakeshott et al. 1982). Whether the activity differences are due to differences in the concentration

or else the kinetic properties of the allozymes remains unresolved (Koehn et al. 1983), as does the exact physiological role (Van Delden 1982) of ADH.

The detailed biochemical studies of these and other polymorphisms were stimulated by the observation of patterns of variation in nature that provided circumstantial evidence of natural selection. It is notable that the most biochemically advanced studies all refer to poikilotherms. The possibilities for environmental selection, particularly with respect to temperature-dependent activity differences of allozymes, would seem to be much greater in poikilotherms than homeotherms. Whilst heterozygosity is generally higher in invertebrate animals than in vertebrates, it is not consistently so. Further, the metabolic activity of homeotherms is very dependent upon environmental temperature, and many species undergo hypothermy or hyperthermy (Kendeigh et al. 1977).

Many potential examples of selection acting on electrophoretically detected protein polymorphisms have been reported in vertebrates. Most of these data come from studies of small mammals and birds. In particular, there is considerable evidence for changes of allele frequency and heterozygosity associated with population cycling in microtine rodents (reviewed by Gaines 1977). These data are compatible with Chitty's (1967) hypothesis that population density changes are due to changes in gene frequency, but direct evidence for such a relationship is lacking. Survival patterns at a number of loci provide indirect support for the operation of natural selection (e.g. Birdsall 1974, Garten 1977) but the data lack repeatability among studies with respect to the effects upon alleles and loci among locations and species (e.g. Tamarin and Krebs 1969).

Several population studies of biochemical variants in birds have also provided some evidence of the operation of natural selection. From studies of laboratory populations, the maintenance of the transferrin (homologous to egg conalbumin) polymorphism in the pigeon Columba livia and ring-necked pheasant Phasianus colchicus has been attributed to heterozygote advantage accruing from greater resistance to infection of heterozygote females' eggs (Frelinger 1971, 1972, 1973, Frelinger and Crow 1973, Lucotte and Kaminski 1976).

Wild birds are particularly amenable to studies of behaviour and reproductive biology and several ecological genetic projects have therefore focussed on comparisons of behaviour and reproductive success among biochemical genotypes. Associations with population density have been reported for the heterozygosity of the Ng locus in the blue grouse Dendragopus obscurus (Redfield 1973, 1974, Zwickel et al. 1977) and an egg albumen locus in red grouse Lagopus l. scoticus (Henderson 1976, 1977). Baker and Fox (1978) found that dominant individuals in flocks of dark-eyed juncos Junco hyemalis were usually heterozygous for L-leucylglycylglycine peptidase, and dominance, which was associated with larger size, was an important component of survival during food restriction. An effect of esterase genotype upon laying date was found in swans, Cygnus olor, together with a suggestion of higher poductivity for esterase and lactate dehydrogenase heterozygotes (Bacon 1979, Birkhead, Bacon and Walter 1983). Evans (1980) reported that breeding success and nestling survival in European starlings, Sturnus vulgaris, was higher for an esterase homozygote genotype, although the clutch size was lower. The results of these studies suggest that ecological genetic studies of biochemical and polymorphisms in birds may be

particularly rewarding.

A few of these studies included the collection of data concerning families. Such data is potentially the most valuable for evaluating the role of different components of any selection that is operating, such as gametic, fecundity or sexual selection. Further, it allows the elucidation of other nonrandom processes of importance to population structure, such as the mating system, inbreeding and assortative mating. The value of family data has been argued forcibly by Christiansen and Frydenburg (1976). Their own data concerning the eelpout Zoarces viviparus was restricted to mother-offspring combinations, and the need to derive paternal information by inference required elaborate statistical models. When paternal data are also available the analysis is much simpler. It is therefore remarkable how relatively few studies have sought family data.

Protein polymorphisms have been used in many studies of the genetic structure of natural populations by employing Wright's F-statistics (see Wright 1979). These studies have in most cases taken the statistical models of neutral theory as their null hypotheses. These hypotheses have not in general been falsified, but hypotheses concerning other possible modes for the maintenance of genetic variation have not been falsified either.

The genetic structure of bird populations, as compared with most others, has had three extra elements of interest. Firstly, the ability to obtain good ecological data concerning standard population parameters, particularly through the use of ringing studies, allows an independent test of stochastic genetic models (see Section 1.4). Secondly, historical information concerning colonisations and introductions provides an independent timescale

against which the predictions of stochastic theory can again be tested (Parkin and Cole 1984b, Parkin 1984). Thirdly, hypotheses can be tested concerning the genetic consequences of the role of bird song in mate choice (e.g. Nottebohm and Selander 1972, Baker et al. 1981, Baker 1983, Zink and Barrowclough 1984, Baker et al. 1984). Most studies have been concerned with the genetic variation found over large geographic areas; this thesis will concentrate on the variation found within populations.

1.3 Aspects of Breeding Systems

The occurrence of particular mating systems in different taxa as assessed by direct observation of behaviour has been well documented (e.g. Wittenberger 1979, 1981), particularly in birds (e.g. Oring 1982) but there has been very little quantification of the proportion of copulations occurring between available individuals of one sex and a particular individual of the other sex. Obviously, even if all copulations were observed, they could not be expected to result in an equal rate of fertilization. Consequently, behavioural observations may not be very informative concerning the more important consequence of mating behaviour i.e. the relative genetic contribution made by each individual to subsequent generations. This kind of quantitative data will ultimately be required to test the adaptive value of different behaviour patterns, particularly where there are marked differences in the behaviour of individuals within populations. Even where the behaviour of individuals suggests that they are monogamous, the reproductive exclusivity of the relationship has only rarely been tested. Also, as pointed out by Halliday (1983) in a review of

studies of mate choice, data on paternity have been obtained only rarely, and 'may be of great importance in providing conclusive evidence of mate choice'.

The occurrence of intraspecific brood parasitism (the depositing of eggs or offspring into the brood of a non-parental adult) has also been receiving more attention recently (e.g. Yom-Tov 1980b). Its detection and quantification involves similar difficulties to those encountered with mating systems. The genetic analysis of putative parent-offspring combinations is also potentially informative regarding this behaviour. If any intraspecific brood parasitism or else cuckoldry occurs then it is desirable, as far as possible, to detect those cases to allow their elimination from genetic analyses of families.

1.3.1 Parentage Analysis Using Genetic Markers

The genetic analysis of families has been used, for example, to exclude or ascribe parentage, to determine the mating system, to detect intraspecific brood parasitism, and to investigate sperm competition (as defined by Parker 1970). Any mendelian genetic marker is potentially informative, and those used include morphological, chromosomal, antigenic and enzyme polymorphisms. For example, Mineau and Cooke (1979) were able to detect intraspecific brood parasitism and extrapair fertilizations (fertilizations occurring between a member of a monogamous pair and a non-mate) in a wild population of Lesser Snow Geese, Anser c. caerulescens, by the appearance of offspring in the clutches of adults with genetically incompatible plumage phenotypes. Plumage varieties have also been used in captive mallards (Anser platyrhynchus) to show that forced copulations (or 'rape' by males

other than mates) sometimes result in successful fertilization (Burns, Cheng and McKinney 1980), and also to successfully investigate sperm competition (Cheng, Burns and McKinney 1983). Eye colour mutants have been used to demonstrate sperm displacement in laboratory-kept Drosophila pseudoobscura (Pruzan-Hotchkiss, Ejianne and Faro 1981, Levine 1982), and body colour mutants have provided data concerning non-random sperm usage and the number of males successfully mating with each female in wild populations of the fish, Xiphophorous maculatus (Borowsky and Kallman 1976). An unusual piece of genetic evidence suggesting extrapair copulations leading to multiple paternity was obtained when two successive clutches produced by a pair consisting of a female tree sparrow Passer montanus and a male house sparrow Passer domesticus each contained both hybrid and apparently true tree sparrow offspring (Cheke 1969).

In wild populations generally, suitable morphological polymorphisms are rare. Consequently, chromosomal and biochemical polymorphisms have usually been used. Chromosomal polymorphisms in the offspring of wild caught inseminated Drosophila pseudoobscra have revealed the frequent occurrence of multiple mating (Anderson 1974, Levine 1982). By using enzyme polymorphisms, multiple fertilization in wild female animals has also been demonstrated in Peromyscus maniculatus (Birdsall and Nash 1973, data reanalyzed by Merritt and Wu 1975), Drosophila melanogaster (Milkman and Zeitler 1974), the isopod Porcellio scaber (Sassaman 1978, cf. Wilson 1981), the snail Cepaea nemoralis (Murray 1964, cf. Wilson 1981), Belding's Ground Squirrel, Spermophilus beldingi (Hanken and Sherman 1981) the polygynous bat Phyllostomus hastatus (McCracken and Bradbury 1977) and the social wasp Polistes metricus (Metcalf and Whitt 1977).

Evidence of multiple inseminations has similarly been obtained in captive populations, including populations of Apis mellifera (Page and Metcalf 1982). In the studies on Belding's Ground Squirrel and Porcellio scaber, sperm mixing was also demonstrated, whilst partitioning of sperm usage occurred in Apis mellifera and sperm competition, detected as a difference in rates of fertilization amongst matings, appeared to occur in Polistes metricus. Experiments on sperm competition in Drosophila melanogaster have suggested that the marker locus used, EST6, may itself have a sperm displacement-release effect (Gilbert, Richmond and Sheehan 1982).

Parentage analysis in captive populations has revealed that dominant males do not sire all progeny in pigtailed monkey Macaca nemestrina and rhesus monkey M. mulatta social groups (Simons and Crawford 1969, Duvall, Bernstein and Gordon 1976, Smith 1980), whereas dominant males in Mus musculus colonies appear to sire offspring exclusively (Singleton and Hay 1983). Similarly, the oldfield mouse Peromyscus polionotus has been shown to have an essentially monogamous mating system where extrapair fertilizations account for about 10% of offspring (Foltz 1981a). Such a high degree of monogamy is believed to be atypical amongst mammalian species (Kleiman 1977).

Many of these studies rely on methods of parentage analysis developed for legal and forensic purposes in humans and domestic animals. An extensive literature has grown concerning the calculation of probabilities of parentage exclusion (e.g. Salmon and Brocteur 1978, Gundel and Reetz 1981, Foltz 1981a), the probabilities connected with parentage attribution (Majumder and Nei 1983) and the reliability of the techniques used (Valentin 1980, Chakraborty and Ryman 1981, Rothman, Neel and Hoppe 1981,

Chakraborty and Ferrel 1982, Lathrop et al. 1983).

Early paternity investigations using blood groups and enzymes on large samples of cattle revealed frequent errors in pedigree records in Britain (Jamieson 1965), and Poland (Ormian 1979). Similar studies on human families have revealed genetic marker mismatches between children and at least one parent, the rate of families containing mismatches ranging from zero for a sample of hundreds of Danish families (cited by Scharfetter 1978), 1.5% and 10% of children respectively representing 5.2% and 39% of families in separate samples from Michigan, U.S.A. (Schacht and Gershowitz 1961), 2.3% of 2839 tested Hawaiian children (Ashton 1980), 1 of 21 pairs of twins tested in the U.K. (cited by Scharfetter 1978), to 13% of 38 Italian families (Hirsch and Vetta 1978) and 30% of children in a sample of "between 200 and 300" women from Southeast England (Phillip 1973). It must be emphasized that these figures may err considerably from the true rate of nonparentage. They have not necessarily been corrected for the non-detection of some cases, which will lead to underestimation, or for a variety of potential sources of error which will lead to overestimation. These factors will be discussed further in Chapter 4.

1.4 The House Sparrow

The house sparrow, Passer domesticus, is now probably the world's most geographically widespread terrestrial bird, occurring endemically in North Africa, and much of Eurasia (Vaurie 1956), and through introductions by man in North and South America, Australasia, South East Asia, South and East Africa and most inhabited islands (Long 1981). Indeed it continues to enlarge its range in many areas, and recently crossed the equator on the western

coast of South America, while its northward movement along both coasts continues rapidly (Smith 1973, 1980). It seems to be an obligate commensal, and appears to have originated in its present form either in the regions of early agriculture in the Nile Valley, North Africa (Summers-Smith 1963), or in the Near East (Johnston and Klitz 1977). The precise taxonomic position of the species remains controversial, but it is currently placed with its congenerics in the family Passeridae (Old World Sparrows), a connecting link between Fringillidae (the true finches) and Ploceidae (weaverbirds) (reviewed by Summers-Smith 1984). The North American sparrows, the Emberizidae, are more distant.

Tegetmeier (1899) and Kendeigh (in Kendeigh 1973) have each described the house sparrow as the 'avian rat', the former because of its pest status and the latter for its value as an experimental animal. The pattern of successful colonisation and expansion makes the analogy even more apposite.

The house sparrow is a granivore, and in most parts of its range its population peaks at the time when cereal crops are ripening. Large flocks assemble in grain fields, and are consequently of some economic importance (Kalmbach 1940, Dawson 1970). Damage to fruit crops has also been reported (e.g. Samuel 1949, Dawson and Bull 1970). Further, the species has been identified as a vector of pathogens (e.g. Gustafson and Moses 1953, Sodhi and Sing 1970, Hubalek 1977), including the occasionally fatal Chagas' disease in South America (Smith 1973). For these reasons, the species generally has official pest status.

House sparrows have been the subject of intensive experimental and observational study, particularly with regard to their physiology, energetics, anatomy and ecology - perhaps more than any

other wild bird. Some of the ecological aspects have been reviewed in Summers-Smith (1963) and Pinowski and Kendeigh (1977). In particular, many data have been gathered concerning the species' breeding ecology in Britain (e.g. Summers-Smith 1963, Seel 1968a, 1968b, 1969, 1970, Dawson 1972, Schifferli 1976), North America (e.g. Weaver 1942, 1943, North 1968, Will 1970, Mitchell et al. 1973, Sappington 1975, Murphy 1978a, 1978b, Anderson 1978, 1979, Lowther 1979a) and elsewhere (e.g. Naik and Mistry 1970, Novotný 1970, Mackowicz et al. 1970, Simwat 1977, Pinowski and Myrcha 1977). In common with other studies of the reproductive ecology of wild birds, these studies were in most cases aimed at understanding the proximate and ultimate factors affecting reproduction. For the most part these factors remain to be elucidated.

Nesting activity in temperate latitudes begins early in the year with the occupation of suitable nesting sites (generally cavities in buildings) by males; some sites will still be occupied from the previous season (after being used as roosts during winter) by pairs or individuals of either sex (Summers-Smith 1963). Sappington (1975) reported that breeding adults surviving to a subsequent year generally (>90%) moved to a new nest site, but this may well have been due to the removal of nests by farmworkers at his study sites during each winter. Summers-Smith found that most pairs were formed following the replacement of a lost mate, and concluded that most of the suitable nesting sites in his study areas were in continuous occupation.

Nestboxes frequently contain nesting material during the winter, and males usually insert some material at the start of their occupation. Nestbuilding proper mostly takes place in the days prior to the first clutch. Complete nests are often built in as

little as 3 days, and egg-laying may begin before completion of a nest (personal observation). The date of the start of egg-laying is negatively correlated with latitude (Dyer et al. 1977), and the proximate cues seem to include daily temperature (Seel 1968b) and photoperiod (e.g. Will 1970 but cf. Murphy 1978). Clutch size is positively correlated with latitude (Dyer et al. 1977). In Britain, the modal clutch size is 4, and the modal number of clutches is 2. The first clutch is generally begun between the middle of April and the start of May, and additional clutches are begun almost as soon as the previous brood fledges (modal incubation period = 12 days, modal nestling period = 18 days). Nesting activity declines rapidly in July, and the last clutches of the year are begun at the beginning of August.

Where buildings containing suitable nest-sites (usually nestboxes in the studies referred to above) are isolated, the breeding population may be considered as forming a series of colonies. Summers-Smith (1954, 1963) reported that the distribution of nests in his study areas (urban and suburban) was not continuous although that of apparently suitable buildings was, and interpreted the pattern as reflecting colonial behaviour. Dawson (1972) found no evidence of a colonial pattern, but was able to show that some locations, particularly those close to farmyards, were preferred. Dawson also found that nesting success (the number of fledglings produced per nest per annum) was lower at locations further from farmyards (detectable at distances exceeding 100m). Mitchell et al. (1973) also reported very significant differences in breeding success between habitats. Farmyards containing livestock provide good feeding sites for sparrows. This is due to the spillage of animal feed, the availability of open feeding troughs

and storage bins and also the preponderance of insects, required for feeding young nestlings, that occur in the vicinity of livestock. Lowther (1979) found a positive relationship between breeding success and the numbers of livestock present at his study farms, and attributed this to the availability of dung fauna.

Schifferli (1978) demonstrated experimentally, by the addition of an extra egg to complete clutches of various sizes, that the fledging success was highest in broods with more than the modal number of eggs. Lack (1947, 1968) has hypothesised that the number of eggs should be maximised at the largest number that can be successfully fledged. That the number laid would appear to be less than this - an observation also made in some other species of nidicolous passerine birds (e.g. Perrins and Moss 1975) - has been attributed to energetic limitations upon the number of eggs that a female can lay (Schifferli 1976), or factors reducing the survival of larger broods after they have left the nest, such as limitations in the ability of the parents to continue to feed them.

There have been some suggestions that older females lay earlier and larger clutches (Seel 1968b, Dawson 1972), but in general the factors affecting differences in clutch size and other components of nesting success between and within females at a particular locality are not well understood.

Both parents contribute to the care of the offspring, but the females' direct contribution to incubation and feeding is significantly greater. Males do not develop a full brood patch, and their role in incubation (restricted to diurnal periods) would appear to consist of preventing heat loss during limited feeding bouts by females. Most feeding of nestlings is, in general, made by females, and the contribution of males varies enormously both during

and among clutches (Seel 1970, Summers-Smith 1963, personal observation). Lone females have been observed to successfully rear some offspring to fledging, but usually fewer than in nests with two active parents (Summers-Smith 1963, personal observation).

Sappington (1975) made an intensive study of helping behaviour at his study site in the Mississippi Valley, North America. He found that both adults and juveniles helped with feeding at most nests, even though they were apparently unrelated to the nest occupants. The total proportion of feeding made by helpers, across all nests, was low (c. 8%), and helpers were not found to contribute to nesting success. Such behaviour is interpreted as misdirected parental care, as also hypothesized for similar behaviour in Geospiza finches (Price et al. 1983). Sparrows have been frequently recorded as feeding the offspring of other species (e.g. Fitch 1949, Hamilton 1952). There have been other studies of the house sparrow both in North America and Britain involving large numbers of colour-marked birds, and none have reported helping behaviour (Summers-Smith 1963, North 1968). Further, helping behaviour should be observable even where birds are not marked as two birds of the same sex should occasionally appear at a nest simultaneously. During this study helping behaviour was recorded on two occasions only, and in each case was restricted to a few visits on a single day. Helping behaviour on a significant scale would therefore appear to be restricted to certain localities.

Most mortality (c. 50-55%) in England occurs prior to fledging and has been attributed principally to starvation (Summers-Smith 1963, Seel 1970, Dawson 1972). Pre-fledging mortality shows a density-dependent effect: percentage mortality increases with clutch size (Seel 1970, Dyer et al. 1977). The population is not known to

fluctuate markedly from year to year, and about 80% of all fledglings do not survive to breed (Summers-Smith 1963). For adults, the period of maximum mortality in Britain is the breeding season. This is unusual among passerines, but has been also found for the starling Sturnus vulgaris and blackbird Turdus merula (Lack 1968: p.300). Breeding house sparrows which have died are occasionally found on nests during the breeding season (personal observation).

The relative importance of different causes of mortality in different populations have not been determined, but predation by raptors (e.g. Schmidt 1972, Yalden 1980) and car accidents (Hodson and Snow 1965) would appear to be important in some areas. In continental populations overwinter mortality may exceed that during the breeding season (Dyer et al. 1977). In a classic North American study of a flock of house sparrows undergoing mortality during a winter snowstorm, Bumpus (1898) hypothesized a posteriori that the body size of house sparrows is subject to stabilizing selection. The interpretation of the data has not been clear-cut, and many reanalyses have been performed (Harris 1911, Calhoun 1947, Grant 1972, Johnston et al. 1972, O'Donald 1973, Lande and Arnold 1983). Grant (1972) and Johnston et al. (1972) concluded that stabilizing selection was operating in females, and directional selection for larger size in males. The optimum female size was considered to be that where the opposing advantages of large body size for counteracting aggression, storing food reserves and reducing the weight to surface area ratio were balanced by the disadvantage of higher required food intake. Lande and Arnold (1983) have argued that the compounding effects of directional selection and of stabilising selection acting upon correlated

characters led to serious overestimates by previous workers of the strength of the selection. They concluded that stabilising selection upon females was strong but 'barely significant'. The results of experimental field studies have been in broad agreement with Bumpus's observations (Rising 1972, Lowther 1977, Johnston and Fleischer 1981, Fleischer and Johnston 1982). Parkin (1984) has critically reviewed these and other evolutionary studies of house sparrows.

Morphological variation in house sparrows has been extensively studied both in introduced (Johnston and Selander 1964, 1971, 1972, 1973, Selander and Johnston 1967, Packard 1967, Lowther 1977, Baker 1980) and endemic (Johnston 1969a, 1969b, 1972, Johnston and Selander 1973) parts of the house sparrow's range. As with parameters of reproduction, very significant correlations have been found between measures of geographical location or climate with components of size and colouration. In introduced populations this variation agrees with Bergmann's, Allen's and Gloger's ecogeographical rules (Mayr 1963). European populations have been found to disagree with Bergmann's rule - more northerly populations are smaller - but agree for Allen's and Gloger's rules (Johnston 1969, Johnston and Selander 1973). This may, however, be a peculiarity of the Oceanic populations sampled, as the reverse trend is found in more continental endemic populations (Pinowski and Myrcha 1977).

The interpopulational variability observed in introduced populations has been attributed to evolutionary changes associated with the colonisation process. Johnston (1973) showed that the covariance among skeletal size variables among European populations is predictable from North American ones, and took this as especially

strong evidence for the action of natural selection upon a genetic component of the size variation. It has not, however, been demonstrated that the observed interpopulational variation in size or reproductive parameters is truly genetic. In the wild bird species studied to date, size variables generally show a high degree of heritability (e.g. Van Noordwijk et al. 1980, Smith and Dhondt 1980, Garnett 1981, Boag 1983), but these data apply only to size variation within populations. Similarly high within population heritabilities have been found for clutch size and laying dates (e.g. Perrins and Jones 1974, Van Noordwijk et al. 1980, 1981, Findlay and Cooke 1982, 1983, Flux and Flux 1982). Environmental differences between populations may well result in a shift in population means (see James 1983), whilst developmental constraints could account for the predictability of covariance among size variables.

In an attempt to quantify genetic differentiation among house sparrow populations, Klitz (1972, 1973, cf. Johnston and Klitz 1977) investigated protein variability in the species using starch gel electrophoresis. Only four polymorphic loci were found, and the level and variability was insufficient to allow firm inferences. Later work by Manwell and Baker (1975) and Cole and Parkin (1981) revealed a higher degree of biochemical variability. Subsequent studies of endemic (Parkin and Cole 1984a) and introduced (Fleischer 1983, Parkin and Cole 1984b) populations have been able to use a larger suite of more variable loci in studies of populational differentiation. These studies have been recently reviewed by Parkin (1984); in general they have found that the observed levels of genetic differentiation are not inconsistent with those expected from stochatic models of gene flow and random genetic drift. The

degree of differentiation is, however, low and other potentially important influences upon biochemical evolution in these populations cannot be excluded.

Gene-flow between house sparrow populations both in Europe and North America is likely to be relatively low for a bird species; ringing studies have shown that movements tend to be limited to those necessary for feeding, and modal dispersal distances prior to first breeding are typically of the order of 1-2 km or even less (Summers-Smith 1963, North 1968, Cheke 1972, Lowther 1979, Fleischer 1983). Some annual migration is known to occur (e.g. Summers-Smith, 1963 Broun 1972), but this appears to be largely restricted to populations occupying areas with more hostile winters.

The house sparrow was considered to be a particularly suitable subject for an ecological genetic study of biochemical polymorphisms for the following reasons:

(i) The species is common and widespread.

(ii) An exceptionally large body of data concerning the species' ecology, behaviour and physiology was available.

(iii) A large degree of electrophoretically detectable genetic variability had already been found.

(iv) Such a study would complement those, previous and in progress, concerning larger scale geographic and temporal differentiation in the house sparrow.

(v) The investigation of many specific aspects was a practical

possibility. Individuals could be trapped, colour-marked and tissue samples taken; access to nests, and hence families and different age-classes, could be gained by providing nestboxes; the high density of nesting allowed many nests to be sampled concurrently, and a reasonably high level of fecundity was expected (Summers-Smith 1963: for England, modal number of fledglings per clutch = 3, modal number of clutches per nest per year = 2). The animal was believed to be large enough to allow tissue sampling to be carried out non-destructively.

(vi) Ringing studies had shown that dispersal, either temporary or permanent, was very limited in English sparrow populations. The occurrence and detection of local genetic processes and differentiation would therefore be facilitated.

(vii) The species has official pest status. Scientific interference with pest species is more readily acceptable for ethical, economic and public health reasons.

1.5 The Study Sites

The choice of study sites was subject to a number of practical considerations. These included accessibility, the availability of locations suitable for fixing large numbers of nestboxes, security of those nestboxes, and the tolerance of people working and living in the neighbourhood towards sparrows, mist-netting and fieldworkers. Two sites were chosen: the Nottingham University School of Agriculture Experimental Farm at Sutton Bonington and the Nottinghamshire County Council Agricultural College Home Farm at Brackenhurst near Southwell. As the house sparrow in Britain can be

regarded as commensal with man, the local agriculture and human demography will be briefly described.

The Sutton Bonington site is situated in the valley of the River Soar approximately 15km SSW of central Nottingham (Ordnance Survey grid reference SK508262), and at an elevation of about 50m. The Brackenhurst site is situated in a more undulating area 18km NE of central Nottingham (grid reference SK698526), but despite being in a relatively exposed location on a small hill, its elevation, at around 60m, is only slightly more than Sutton Bonington's. The locations of the two sites, and their environs, are shown in Figures 1.1 - 1.3.

The entire area containing the study sites is occupied by house sparrows and nesting takes place in all areas of human habitation: i.e. towns, villages, hamlets and farms. The agricultural practices at both farms are fairly similar, and typical of the region, involving a mixture of livestock, grass, and winter and spring sown cereals. There were many more cows and pigs, and a greater area of grass pasture, in the immediate vicinity of the nestboxes at Brackenhurst, but there were a number of sheep near to the Sutton Bonington site, an animal absent from the vicinity of the Brackenhurst site. The nestboxes at both sites were within a few hundred metres of housing, student accommodation and gardens. Both were near to larger areas of human habitation. The Sutton Bonington site is less than 1km from the village of Sutton Bonington itself, and 2km from the small town of Kegworth. Brackenhurst is 1km south of Southwell (population 5129), a slightly larger town than Kegworth (population 2814) (Anon 1977).

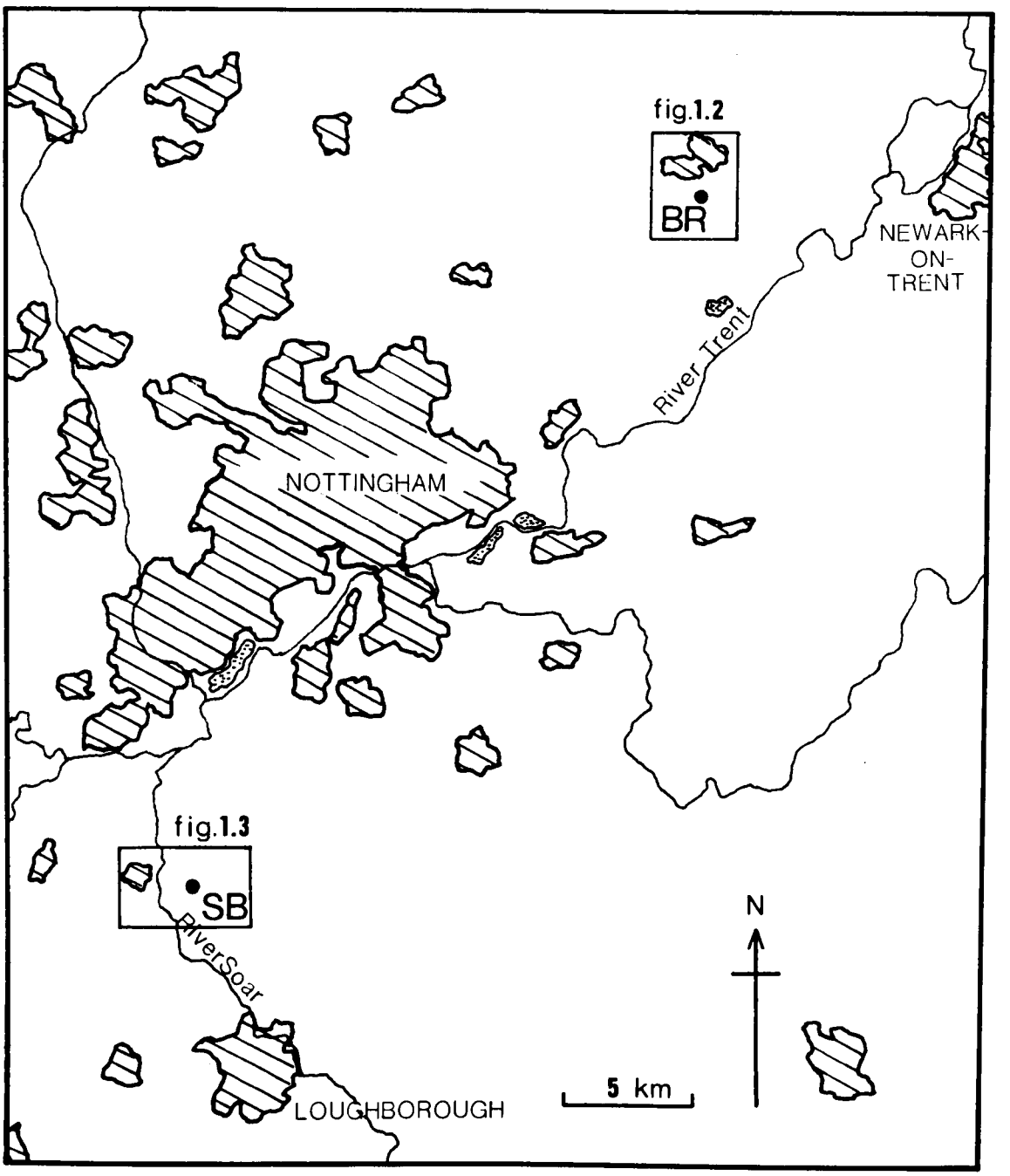


Figure 1.1

Map showing locations of study sites. cf. Figures 1.2 and 1.3.

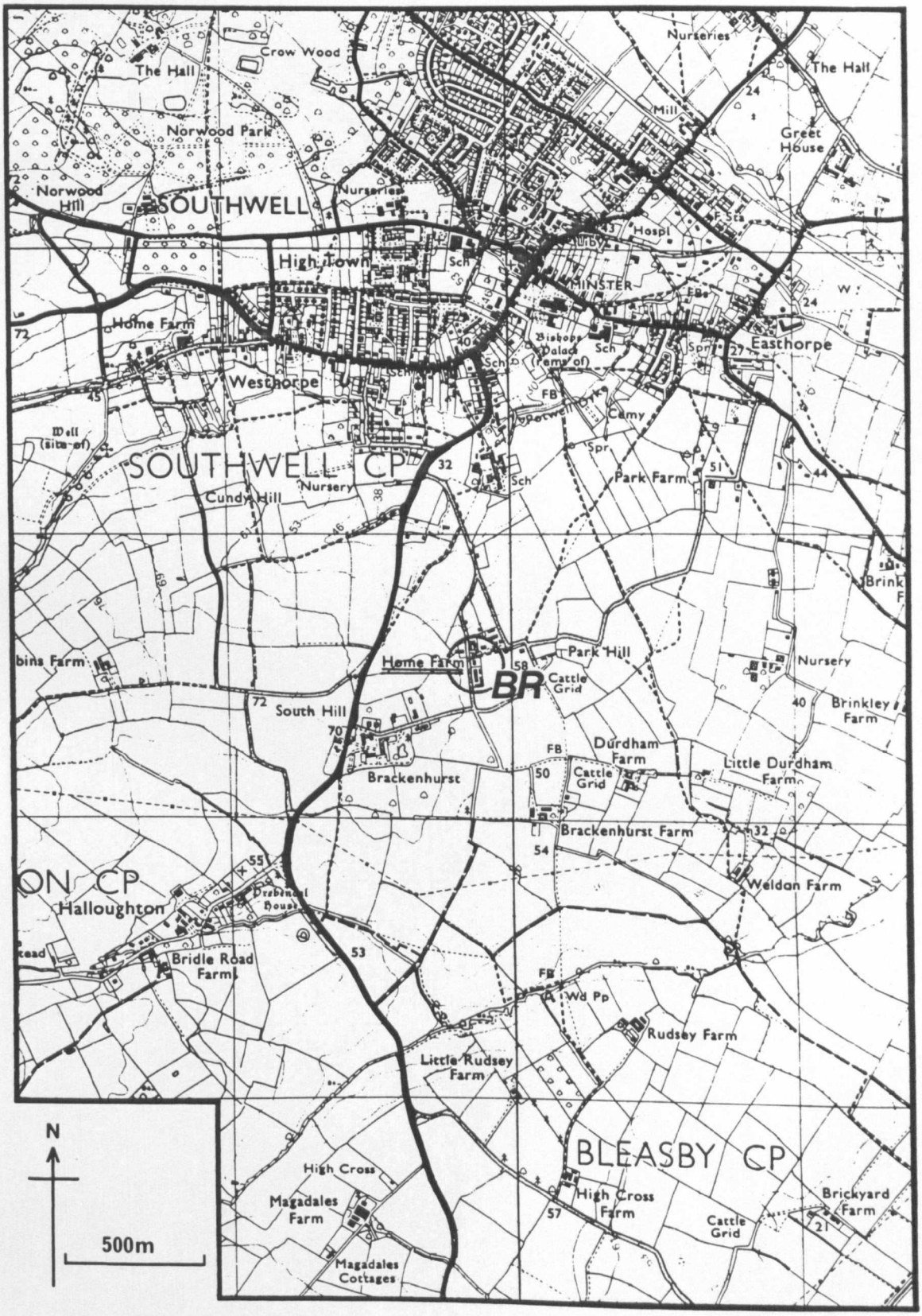


Figure 1.2

Extract of Ordnance Survey map containing Brackenhurst (BR) Home Farm study site (circled). Spot-heights and contours are shown in metres.

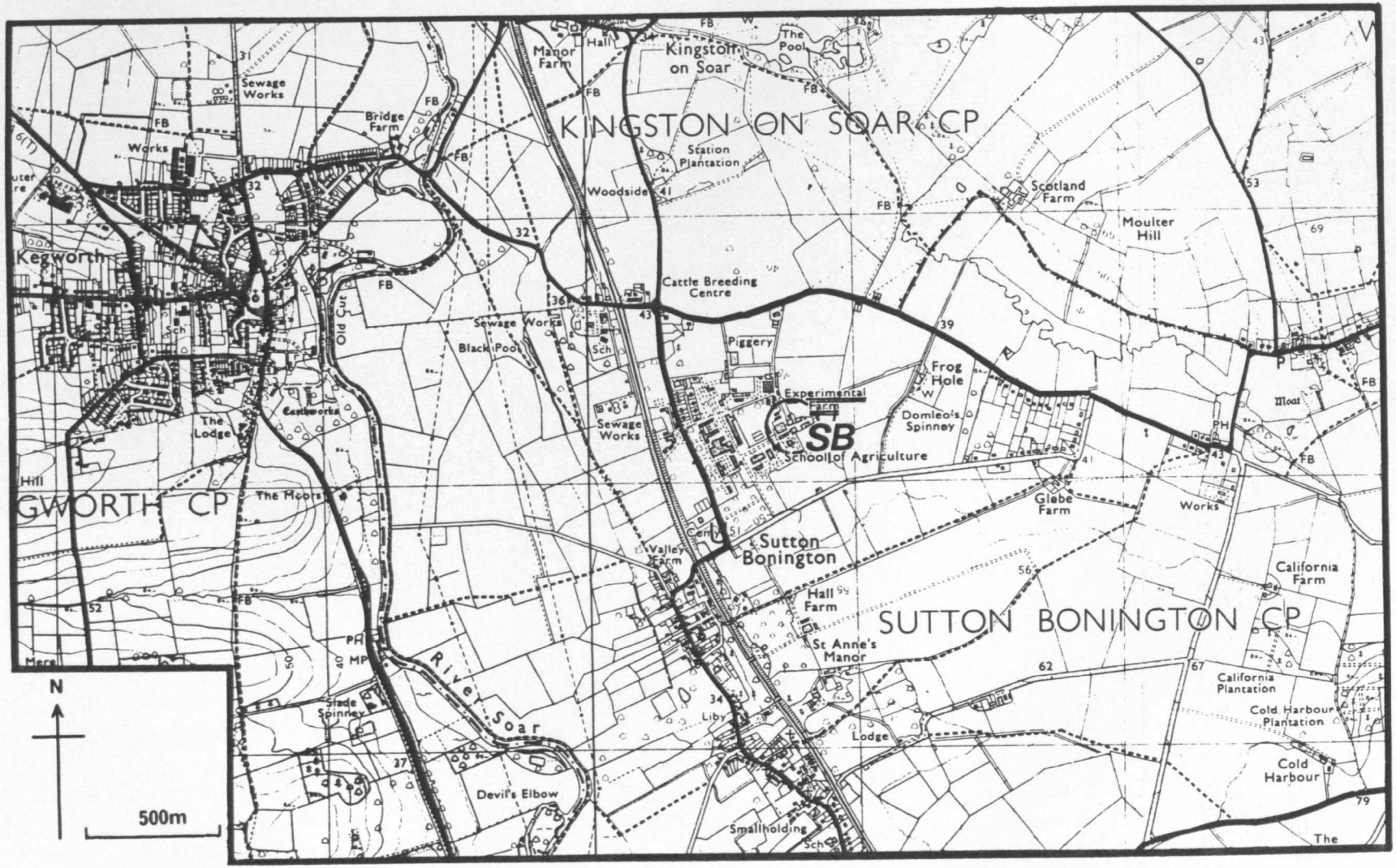


Figure 1.3 Extract of Ordnance Survey map containing Sutton Bonington (SB) Experimental Farm study site (circled). Spot heights and contours are shown in metres.

1.6 Aims of the Study

The principal aims of the study were as follows:

(i) To investigate in detail the genetics of as many marker loci as possible, without destructive sampling (Chapter 3).

(ii) To use the marker loci to investigate the possibility of extrapair copulations and/or intraspecific brood parasitism (Chapter 4).

(iii) To investigate in detail the distribution of alleles and genotypes among age, year and sex classes and test for the possible operation of nonrandom processes. The design was to include reasonably sized samples from two populations in at least two years (Chapter 5).

(iv) To investigate the possibility of nonrandom forces acting on the genotypic distribution of mate pairs (Chapter 6).

CHAPTER 2

METHODS

2.1 Field Methods

2.1.1 Provision of Nestboxes

The requirements of a nestbox were that it should provide an attractive, adequate and secure nesting environment, that it should be easily and cheaply produced, and allow easy access for the fieldworker. The internal dimensions were close to those used in other studies (e.g. Dawson 1972). All the boxes used in this study were of the same dimensions and design; it has been shown that the breeding success of several species is affected by the size and construction of nestboxes (Moeed and Dawson 1979, Karlsson and Nilsson 1977). The exact design is shown and described in detail in Figure 2.1. The boxes were fixed with a dowel perch, enabling colour rings to be seen. All the boxes were treated annually with creosote wood preservative.

Locations for nestboxes were selected with regard to accessibility (using a ladder extending to 5.2m), availability of a good vantage point for the observation of parent birds, reduced exposure to weather, possibilities for mist-netting, and the proximity of naturally occurring nests. The numbers of nestboxes available and used for nesting in each season are summarised in Table 2.1. The number of boxes used, and the number of clutch starts in each increased during each year at each site. Site plans and precise nestbox positions are given in Figures 2.2 and 2.3. A typical farmyard view showing nestboxes is shown in Figure 2.5.

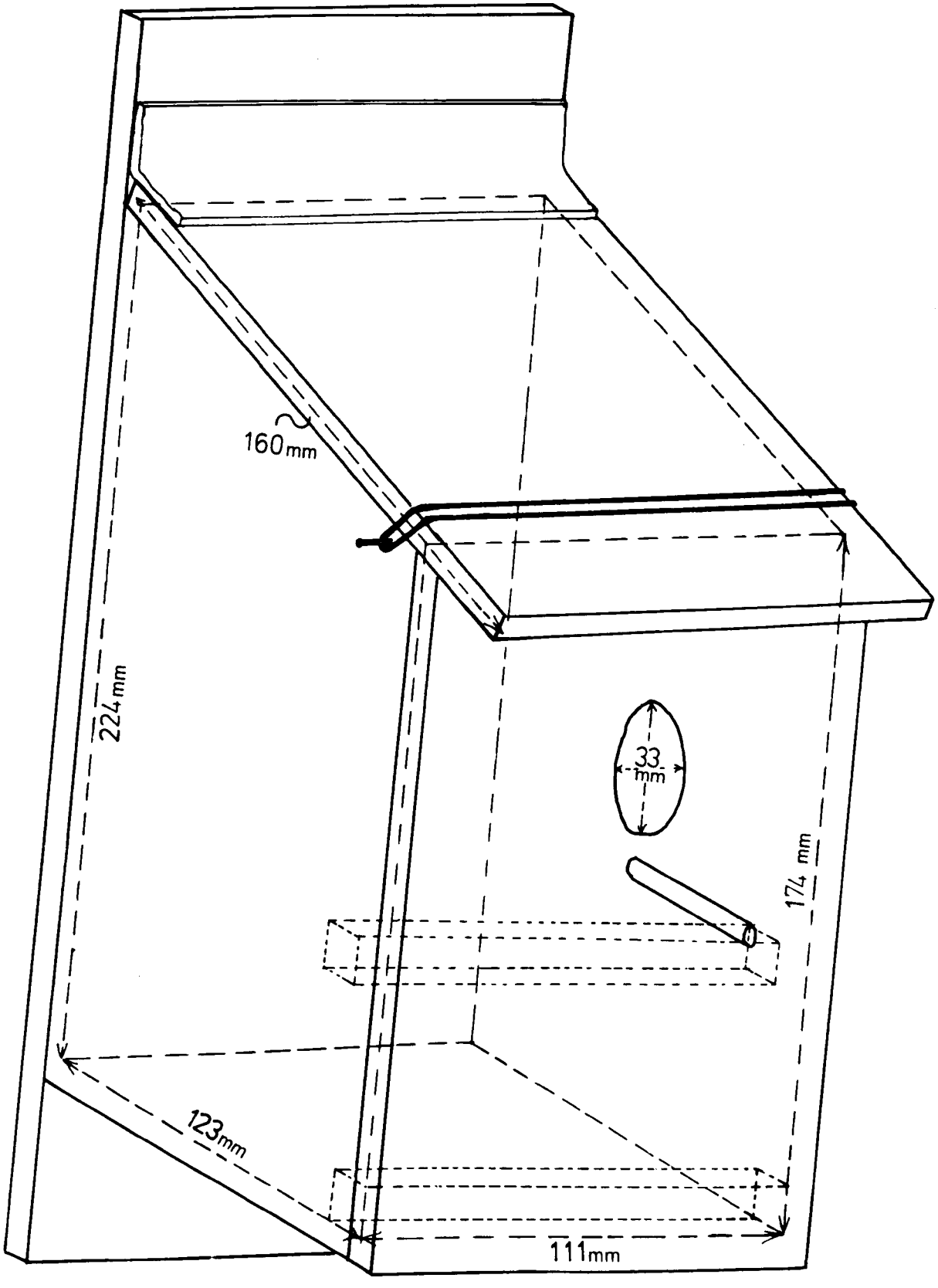
Most of the nestboxes provided during the first winter (1979-1980) were placed in covered areas. A group of 24 in the cowshed

Figure 2.1

Nest box design

The boxes were constructed from 6mm exterior grade plywood; 19 could be made from one 120 x 240 cm sheet. The component plywood panels were joined using fine panel pins, together with a waterproof wood glue for fixing the main part of the box to the back board. 15mm x 15mm ramin beading was used to provide extra strength along the front bottom join, and another piece positioned so as to provide support for traps (Figure 2.4). The box lids were initially hinged using carpet binding adhesive tape, but this was replaced after two years by stapled strips of 40mm wide elasticated seat webbing. The lids were held in position by an elastic band stretched between a nail or small screw on each side. A 40mm length of 6mm dowel was attached perpendicular to the front, 10mm below the entrance. The internal dimensions are indicated.

/contd.


160mm
224mm
33 mm
174 mm
123mm
111mm

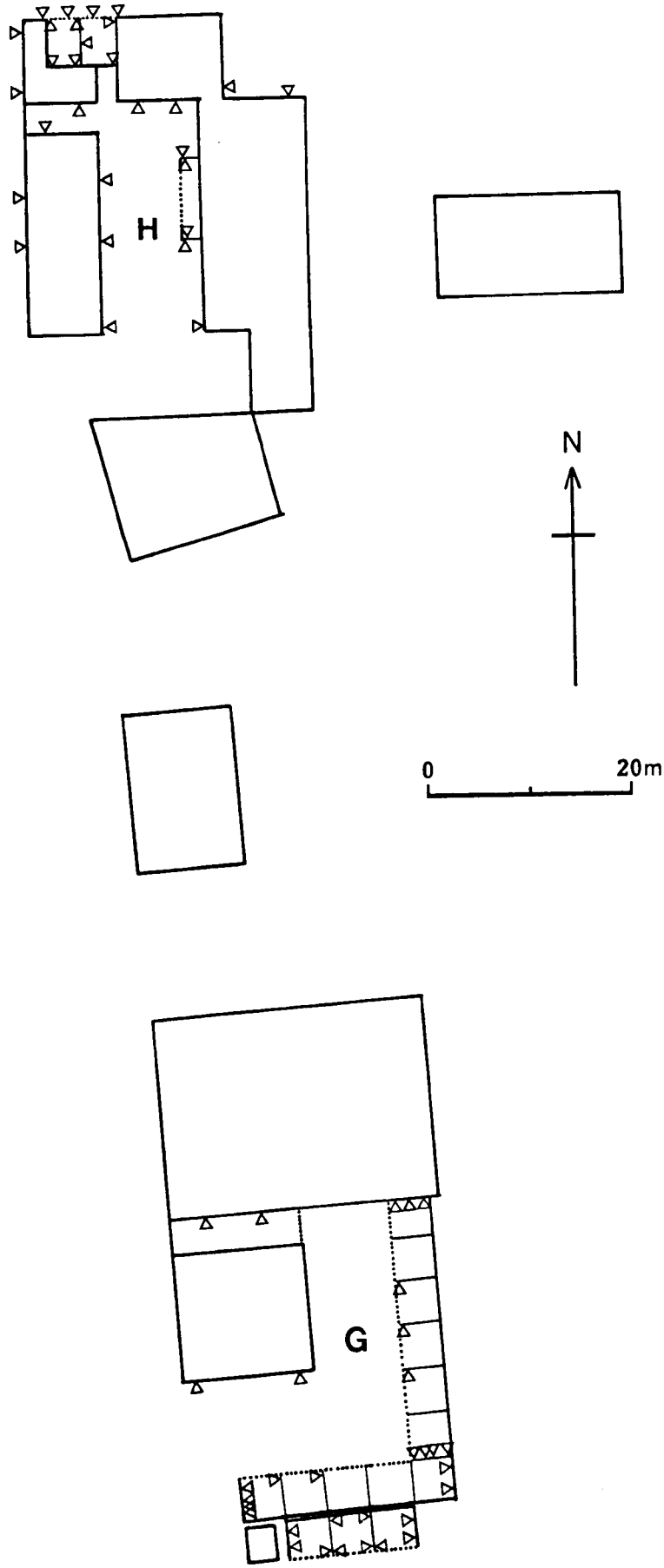


Figure 2.2

Ground plan of buildings at Brackenhurst (BR) site showing positions of nestboxes in 1982. The two parts of the site were designated 'G' and 'H'.

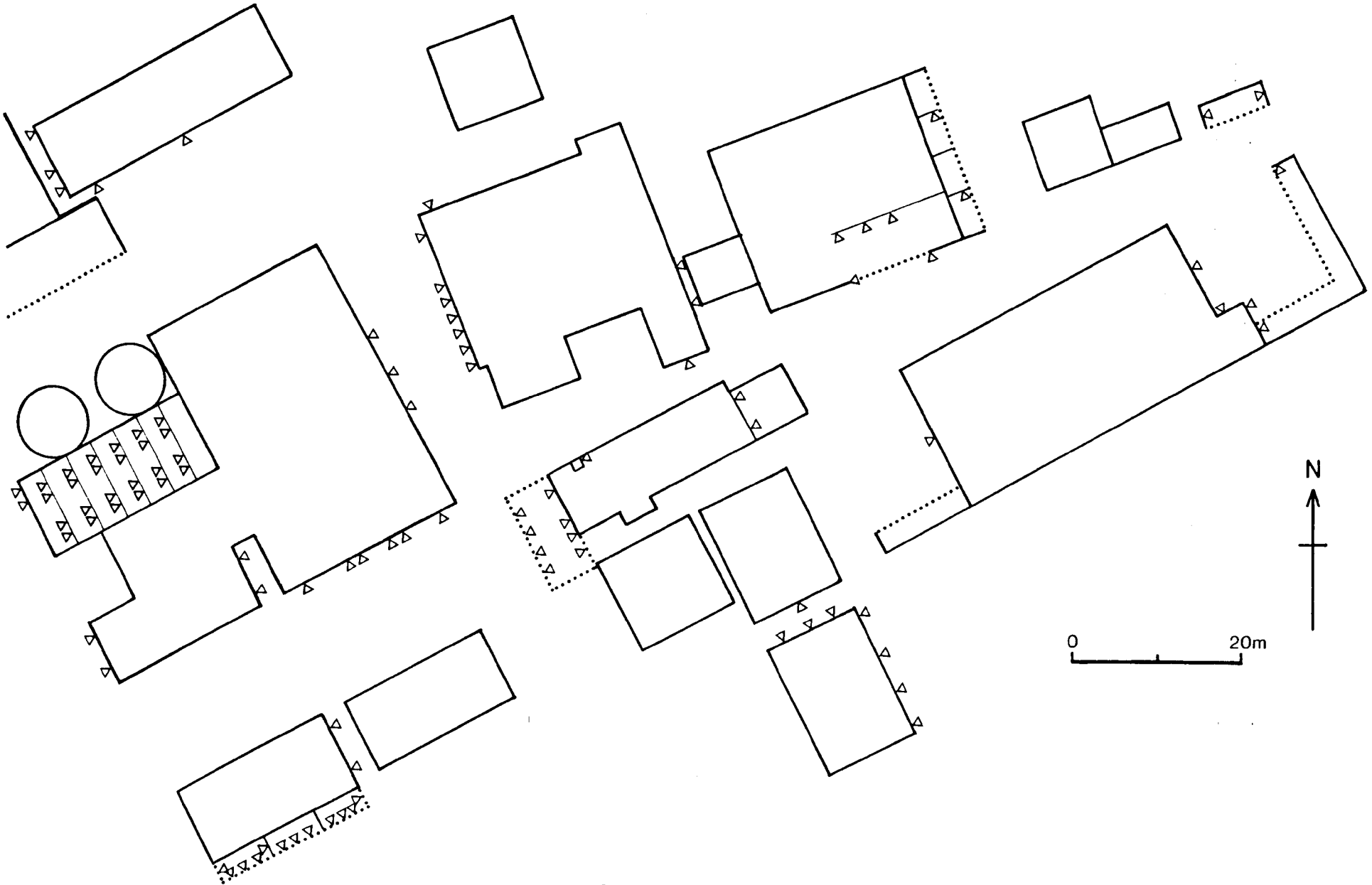


Figure 2.3

Ground plan of buildings at Sutton Bonington (SB) site showing positions of nestboxes in 1982.

at Sutton Bonington replaced boxes provided in a trial study by D.T. Parkin and S.R. Cole during the previous year. Most boxes were put out during the winter months December - March, but occasionally extra boxes were provided as late as the end of May if a particular part of a site was proving to be a preferred nesting area. Those nestboxes provided for 1980 on exterior walls were used only to a very limited extent, but more success was had with additional exterior boxes provided in subsequent years. Where boxes were placed near to naturally occurring nests the latter were, if possible, removed. By persistent nest removal many birds were persuaded to use a nestbox instead, but where the nests were inaccessible the provision of nestboxes was not always productive.

No attempt was made to orientate boxes in any particular direction, and there was no obvious effect of orientation upon subsequent useage. Boxes less than 3m above an adjacent surface, whether the ground or the roof of a lower building, were only rarely used.

Old nests were removed from all nestboxes before the start of each new nesting season. This was in part intended to restore the available space inside the box, as new nests are often otherwise built over old compressed material, and also to reduce the flea content.

2.1.2 Monitoring Nestboxes

Regular visits to each nestbox, with the purpose of monitoring the performance of individual adults and nestlings, began at the start of April each year. The two sites were visited on alternate days throughout the breeding season. The earlier visits in each year were made to all boxes at maximum intervals of 4 days (usually

Table 2.1

Nestbox availability and use at each site during the 3 years of the study.

Year	Site					
	SB			BR		
	Available	Used	Total clutch starts	Available	Used	Total clutch starts
1980	54	20 (37.0%)	29	35	14 (40.0%)	26
1981	103	61 (59.2%)	99	64	41 (64.1%)	79
1982	103	75 (72.8%)	138	60	49 (81.7%)	99

every 2 days in 1980), but those showing no nesting activity were subsequently checked at more irregular intervals. In 1981 and 1982, newly detected clutches were not reexamined for a further 4 days, and in all years complete clutches were checked for the presence of hatchlings about 10 - 11 days after the completion of a clutch. For this purpose the date of clutch completion was calculated on the basis of one egg being laid per day (Summers-Smith 1963, Seel 1968a).

Each individual within a nest was marked on the first visit after hatching by clipping a different claw. On this and subsequent visits (usually made at approximate 48h intervals) all nestlings were weighed to 0.1g using a 50g Pesola spring balance. The accuracy of the balances was checked regularly. Where nestlings were particularly small at the time of marking the appropriate claws were re-clipped on a later visit.

2.1.3 Sampling Adults and Parents

Adult birds were trapped regularly away from the nestboxes. This was to allow the estimation of genetic and demographic parameters, and, by marking all trapped adults with colour-rings (see Section 2.1.4), enable their identification during subsequent nesting attempts.

The trapping was achieved by the frequent use of mist nets during spring and summer at both sites. Most of the birds caught in this way were either feeding on animal feed supplied in troughs in the open (at Brackenhurst) or on stored cereals and feed in open grain bins inside a mill (at Sutton Bonington). About one-quarter of identified parents were originally caught by this means (Table 2.2). Additionally, mist nets were occasionally deployed across

Table 2.2

Trapping techniques used for the original capture and marking of birds subsequently identified as parents of successful clutches (both sites).

Method of initial capture	Parents of successful clutches	
	Females	Males
Random mist-netting	27	43
Mist-netting in front of nestboxes	38	36
Total from mist-netting	65	79
Trapping in nestboxes	68	54

the fronts of open sheds containing nestboxes, thus ensuring the inclusion amongst captures of a high proportion of parents of clutches under study. This accounted for a further quarter of identified parents (Table 2.2). Nets were not used in this way for more than one hour in any particular position.

Parents were identified by observing their activities at nestboxes, usually by using 10 x 50 binoculars or a tripod-mounted 20 x 60 telescope from the inside of a conveniently parked vehicle. Reading colour-rings was often a time-consuming procedure, requiring the observation of many nest-visits by each parent as sparrows' tarsi are frequently obscured due to their squatting posture. Where a parent was not already marked, attempts were made to trap it inside the nestbox by using a spring-loaded trap positioned just behind the nestbox entrance (Figure 2.4). The trap design was based on one kindly supplied by Dr. P.G.H. Evans (Oxford University). The trap was more effective when the spring and door, which were painted black, were disguised by nesting material such as feathers. Trapped birds were removed from the nestbox using a net designed to enclose both the box lid and the operator's wrist. To reduce the possibility of desertion by parents trapped in this manner, and to minimise any loss of data concerning nestlings irrespective of their parents, these traps were only used, subsequent to a trial period at the beginning of the 1980 season, at the time of ringing and bleeding of pulli (see next section). In these circumstances a trap was set whilst the nestlings were temporarily removed from the nest. If trapping attempts are postponed in this way the nestlings will be large and therefore have to be removed as they would otherwise trigger the trap mechanism. Also, parents only occasionally enter the nest when it contains large nestlings as most

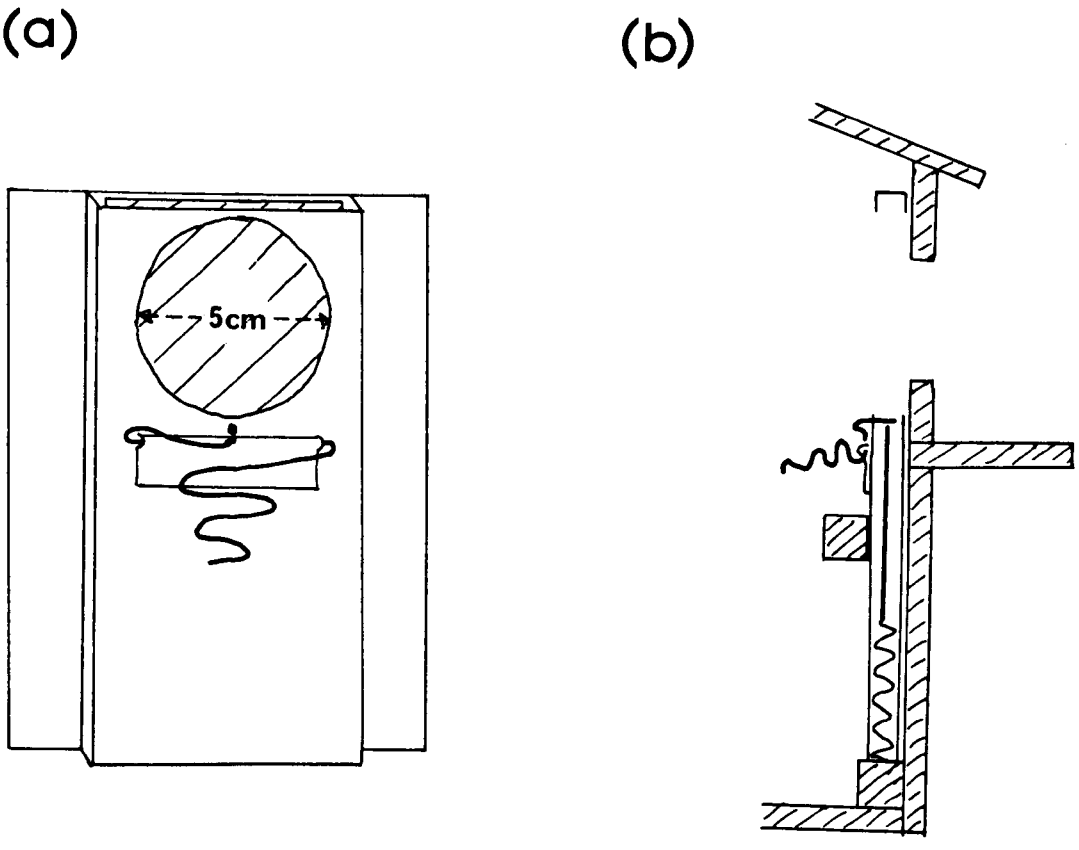


Figure 2.4

Springloaded nestbox trap. The trap was painted black to make it less conspicuous. The hole diameter was larger than the nestbox entrance (cf. Figure 2.1).

(a) Rear view of trap in closed position.

(b) Diagram of trap set in nestbox. The spring trip was disguised with nest material (feathers and straw).

feeding and the collection of faecal sacs takes place from the perch (Figure 2.6). The use of nestbox traps was about equally important to the study as the use of mistnets. Relatively fewer male than female parents were trapped in nestboxes (Table 2.2).

2.1.4 Processing Birds in the Field

The trapping, handling and ringing of birds used in this study was carried out in accordance with the requirements of a Nature Conservancy Council permit. The advice of the British Trust for Ornithology (BTO), as outlined in its Ringer's Manual (Spencer, 1976), was closely followed. As wild birds in general, and house sparrows in particular, are known to transmit disease to humans (e.g. salmonella: Macdonald and Brown 1974, Cornelius 1969; encephalitis: Lord et al. 1974) hygenic precautions, particularly with regard to hand-washing and the antiseptic protection of wounds, were observed at all times.

All birds trapped, and all nestlings prior to fledging, were measured and tagged and had a blood sample taken. Nestlings were processed between 10 and 13 days after hatching. The age and sex of all birds was recorded wherever possible. Adults are very obviously dimorphic for plumage patterning and colouration, but the sexes of nestlings and younger juveniles could not be distinguished. This species is slightly unusual in that a full moult immediately follows the breeding season in both adult and juvenile (first-year) birds so that juveniles attain full adult plumage and cannot be distinguished from freshly moulted, older birds (Svensson 1975, Newton 1966). The incompletely ossified skull of house sparrows of up to about 200 days of age can be used as a distinguishing feature (Svensson 1975, Nero 1951, Niles 1973) but the extra handling time

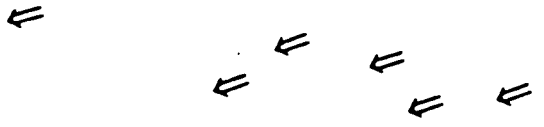


Figure 2.5 View of farmyard buildings at Sutton Bonington, illustrating typical nestbox positions. Six nestboxes (arrowed) are in view.

Figure 2.6 Adult male feeding 16 day old pullus from the perch at a Brackenhurst nestbox. The nestbox is in a covered area.

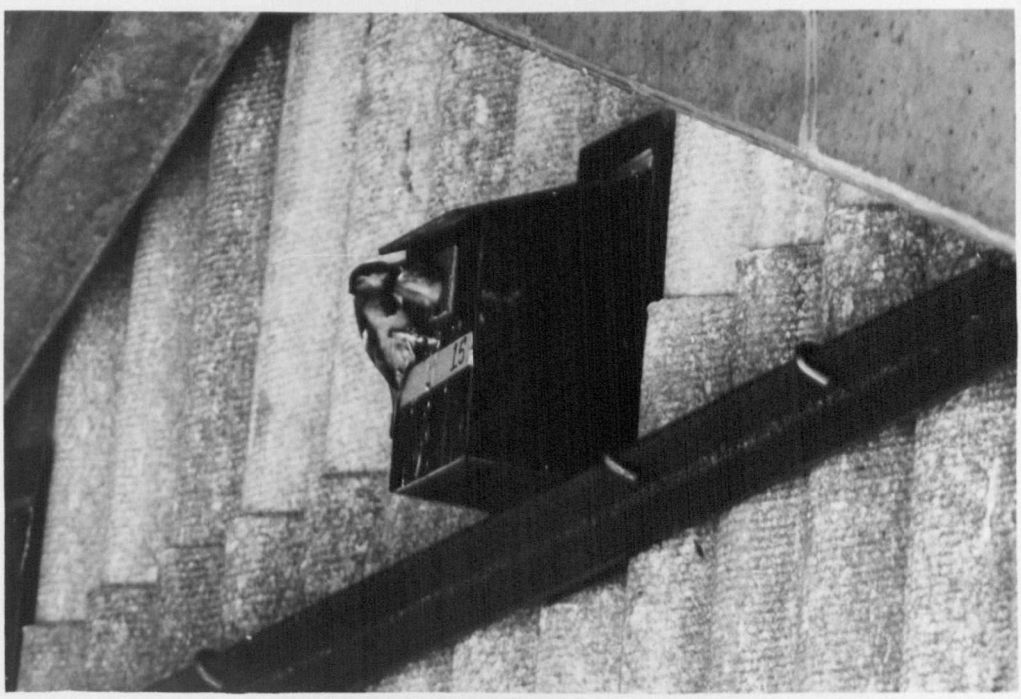

15

Figure 2.5 View of farmyard buildings at Sutton Bonington, illustrating typical nestbox positions. Six nestboxes (arrowed) are in view.

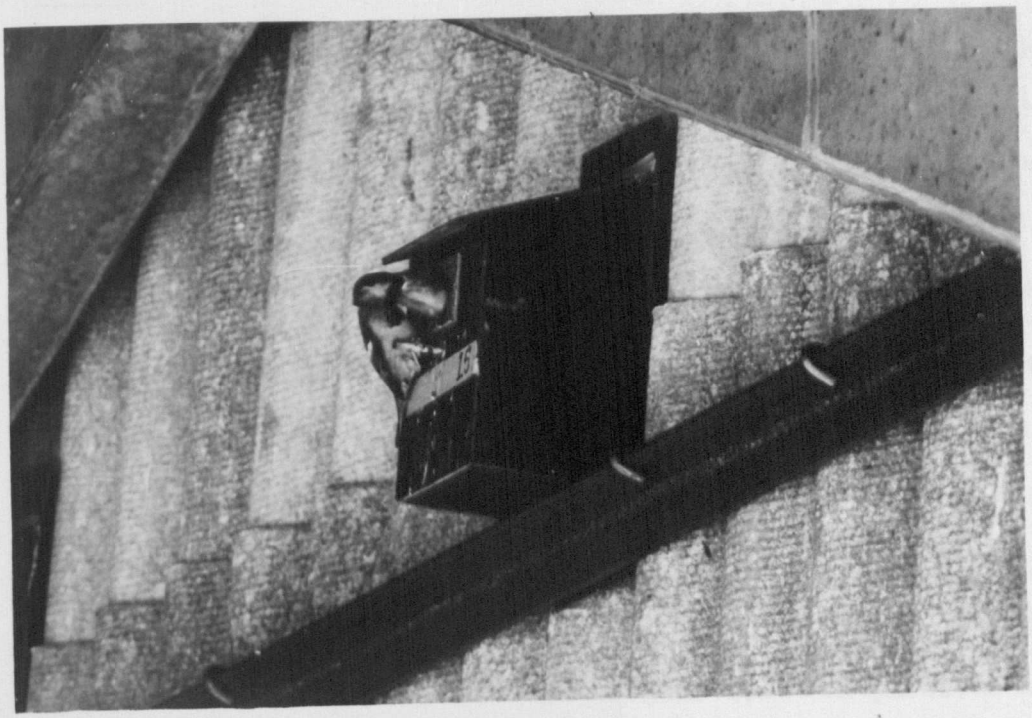


Figure 2.6 Adult male feeding 16 day old pullus from the perch at a Brackenhurst nestbox. The nestbox is in a covered area.

required for this was not considered worthwhile for the small amount of extra information that would have been obtained.

Initially, juveniles resemble adult females but they can be distinguished until their moult is complete by feather condition and the shape of the outermost (10th) primary wing feather (Cheke 1967). Although some discrimination between the sexes is possible earlier (Cheke 1967, Harrison 1961), juveniles cannot be sexed reliably until the onset of the post-juvenile moult. The small proportion of adults and juveniles caught in late summer and autumn were checked for stage of moult and consequently those fully moulted birds that could not be aged (<6% of all captures) were distinguished from those known to be either less than, or else greater than, one year old.

Metrics

Adults and juveniles were measured for weight and wing, tail and tarsus lengths, and nestlings were measured for weight and tarsus length. Weight was measured using a 50g Pesola balance, interpolating to the nearest 0.1g. Wing length was measured to the nearest 1mm by the method of maximum chord (Svensson 1975) by pressing the underside of the folded right wing, flattened and straightened by the thumb of the hand holding the bird, against a stopped end wing rule. Tail length was determined to an accuracy of 1mm by sliding a thin rule between the tail feathers and under-tail coverts until it came to a stop at the root of the tail (Svensson 1975), whilst holding the tail feathers parallel against the surface of the rule. Where the normally measured longest wing or tail feathers were lost, damaged, or in moult, this was noted and the measurement was not used in any subsequent analysis. Tarsus

length was measured to an accuracy of 0.1mm using a good quality sliding vernier caliper with a thumb-operated lock release.

Three different tarsus measurements were used: TAR1, TAR2 and TAR3. TAR2 was the distance from the notch at the back of the intertarsal (mesotarsal or tibiotarsal) joint to the anterior distal end of the lowest undivided scute. This is the measure most commonly used by other workers (e.g. Smith and Zack 1979). TAR1 was the distance from the nuchal notch to the lower end of the 5th scute (counting up the middle toe) from the division of the two outermost toes. In almost all cases this scale could be identified as the first one immediately above the foot not to articulate when the tarsus was viewed in lateral profile and the toes moved. TAR3 was the measurement from the nuchal notch to the lower end of the scute immediately distal to that used to determine TAR1. TAR1 and TAR3 are illustrated in Figure 2.7. In 91% of cases TAR3 = TAR2, but in the remainder TAR1 = TAR2 due to the occasional division of the lower scute. TAR1 was felt to be the most valuable measurement since its measurement did not rely on either the apparent variability of scale division that occurs between birds or the posture of the foot, and it was used from the start of this study (but see Section 6.3.4). TAR2 and TAR3 were also measured commencing in August, 1981.

Ringing

All nestlings and adults were tagged using uniquely numbered metal rings supplied by the British Trust for Ornithology (BTO). Adults and juveniles were also individually marked using a unique combination of three celluloid colour-rings (supplied by A.C. Hughes Ltd.). All marked birds, except for 8 with one missing or deformed

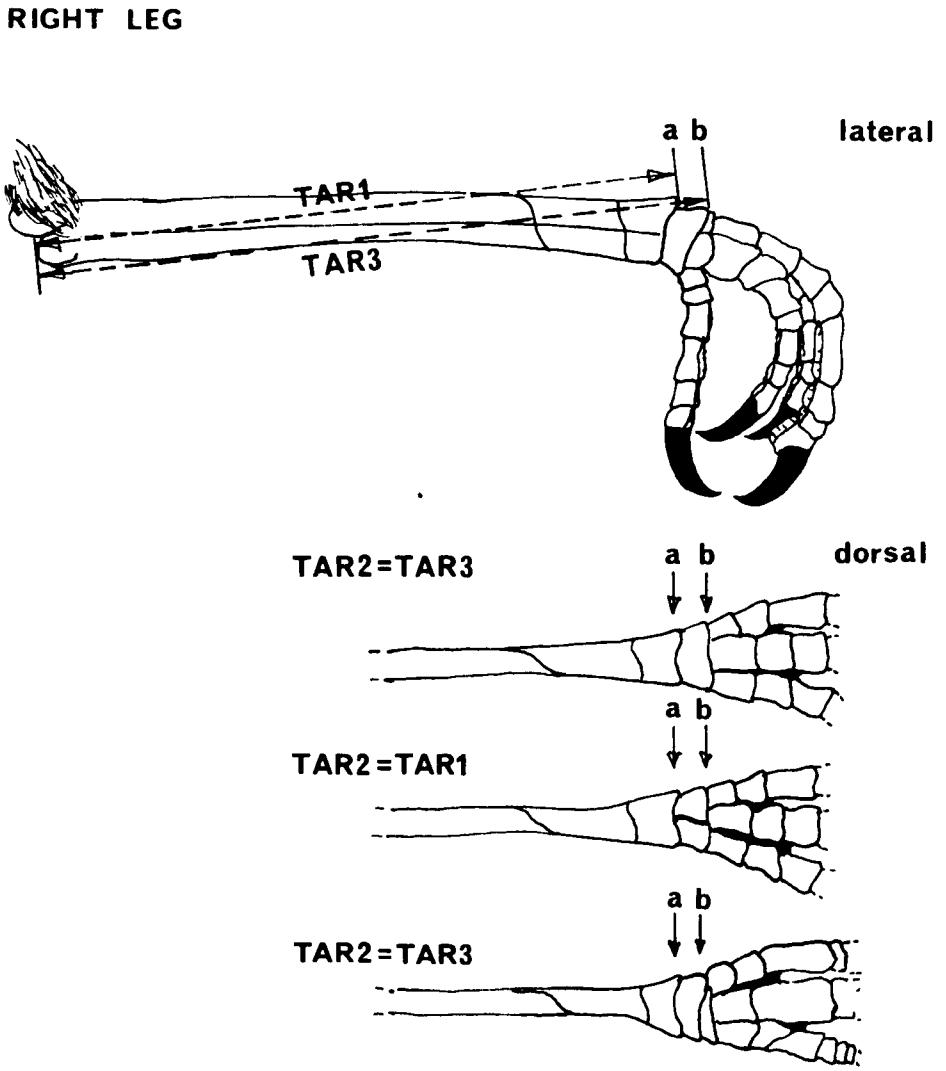


Figure 2.7

Tarsus measurements as taken with sliding vernier calipers (refer to Section 2.1.4). The lower edges of the scales used for measurements TAR1 and TAR3 are labelled a and b respectively. Dorsal views of these scales in three different individuals are included.

tarsus, carried two rings on each leg; combinations requiring fewer rings were not used to avoid the possibility of confusion due to ring loss. Colour-rings were lost occasionally and so attempts to minimise this were made in 1982 by sealing them with acetone applied to the ring split by a fine brush. Eight colours were used: white, mauve, red, black, light blue, yellow, orange and light green. Ring combinations were recorded as a sequence going up the right leg then down the left. With 4 positions available for the BTO ring the 8 colours provided a potential $8^3 \times 4 = 2048$ combinations, each of which could be used on birds of both sexes. As there were many more combinations available than required the colours which proved to be most readily distinguishable - red, orange, yellow and white - were used most often. No colour code was used at both sites, though the probability of a bird migrating between the two sites is known to be extremely low. When observing colour combinations care was taken against colours being misidentified due to poor light or dirt on rings. For example, it was found that in poor light the colours light green and light blue could be confused when viewed through the telescope; the possibility of errors was avoided by checking that alternative colour codes had not been used.

Tissue Samples

As outlined in Chapter 1, there was a requirement to score each individual, without harm, for a suite of genetic loci. The most satisfactory means of achieving this was found to be by taking a blood sample and subsequently analysing it electrophoretically for a series of polymorphic protein loci (see Section 2.2). This approach has also been used in studies involving live birds by Bacon (1979), Whitehouse (1979) and Evans (1980). The use of muscle or

feathers as a tissue source was also considered. Pectoral muscle biopsy undertaken as by Baker and Fox (1978) and M.C. Baker (1981 and personal communication) was assessed using a group of birds kept in the laboratory, but muscle in this species proved to be no more useful a tissue than blood (see Section 3.1.4 and Table 3.1), and its removal was felt to be more traumatic and less convenient. Though variability in feather proteins between species has been used as a taxonomic tool, these proteins are remarkably monomorphic within species (e.g. Brush 1976; Knox 1980).

Approximately 0.5ml blood was taken from the right jugular vein of each bird as described by McClure and Cedeno (1955). Birds have asymmetric jugular veins and that on the right is the larger. It was found most convenient to hold the bird in the left hand with its neck between the first and second fingers and the right wing kept folded by gentle pressure exerted by the thumb at the carpal joint. Gentle pressure placed by the side of the thumb upon the neck could be used to restrict blood flow from the vein and thus make it more prominent. A small area of feathers and skin on the appropriate part of the dorsal surface of the neck was wetted slightly with ethanol and the sample taken using a 2ml disposable syringe fitted with a 25G x 5/8 disposable hypodermic needle. The less travel required by the plunger of the 2ml syringe make it easier to use than the longer 1ml model. To prevent the sample from clotting the syringe and needle were heparinised beforehand by expelling any surplus air, taking in about 0.1ml heparin solution (Weddel Pharmaceuticals Ltd.: Heparin sodium, 5000 I.U./ml), then expelling this back into the still inverted bottle. This left a small bead of heparin (c. 0.01ml) in the syringe. On removal of the hypodermic from the jugular vein bleeding was prevented by quickly

pressing a paper tissue swab on the skin surface and holding it in position with the thumb for about 30s. The needle was removed from the syringe, and the contents expelled into a 2ml screw-top Nunc tube (Nunc 3-66656, 43 x 12.5mm). The tube was shaken to ensure thorough mixing of the heparin and then placed in an insulated cold-box at c. 4°C.

After being bled, birds were kept for a few minutes in a cotton bag or holding box. No immediate mortality was observed in the field or over several days amongst 20 birds kept in the laboratory and bled by this procedure. Alternative methods were investigated initially. Cardiac puncture (Whitehouse 1979) was tested using caged sparrows but was found to be difficult to use on a small species and occasional mortalities resulted. Brachial venipuncture (Evans 1980) was used at the start of the study but was much slower than jugular bleeding and the consequent clotting often resulted in small or partly lysed blood samples.

At 0.5ml the volume of blood taken provided an easily handled quantity and was within estimated safety limits. The blood volume of flying birds is about 7 to 9 % of body weight (Kovách et al. 1969); it will therefore exceed 2ml in the average house sparrow and equal about 1.4ml in the smallest bled nestling. Kovách et al. (1969) found that birds are much more tolerant of blood loss than are mammals and that about 50% of the total volume can be taken.

Blood samples were separated into packed cell and plasma fractions on the same day as their collection by centrifuging for 10 minutes at 1500 x g. Erythrocytes were washed once by mixing well with 2 volumes of Ringer's solution or 0.9% saline and recentrifuging. Both fractions were stored at -80°C.

Retraps

All recaptured birds were remeasured and repeat blood samples were occasionally taken. Ringed pulli found to be still present on a subsequent visit to the nest were remeasured.

2.2 Laboratory Methods

2.2.1 Preparation and Storage of Blood Samples

At the end of the field season, blood samples were prepared for subsequent electrophoretic analysis. Although for mammalian studies the simple lysis of erythrocytes in a hypotonic solution is adequate to provide a solution of proteins for electrophoresis, the nucleated nature of avian cells results in their requiring a more extensive treatment (Sibley et al. 1974, Whitehouse 1979). Avian red cells are generally more difficult to lyse, and even following only partial lysis (such as occurs in freezing and thawing), they form a thick, electrophoretically useless gelatinous mess. The most satisfactory results for house sparrow erithrocytes were achieved by adding one volume of 0.1% Triton X-100 (Sigma T-6878) to the thawed sample, followed by thorough vortexing to separate and lyse the cells and then by high speed centrifugation at 15000 x g for 20 minutes at 4°C. The whole of this procedure was conducted in the same 2ml Nunc tube in which the sample was originally collected. It was found to be unnecessary to remove the supernatant to a new tube before refreezing at -80°C as the pellet in the base of the tube did not noticably resuspend even if the tube was thawed and refrozen several times. Samples were centrifuged in pairs, one stacked above the other; they were separated by a wad of tissue paper and the rounded base of the centrifuge holder was also padded

to prevent tube damage. Rejected preparative techniques included sonication, homogenisation, and different triton concentrations and amounts of centrifugation.

To minimise refreezing and thawing effects and for general convenience during electrophoresis, 30µl drops of plasma and prepared erythrocyte lysates were transferred to the individual wells of flat-well microtiter plates (Sterilin M29A). At least 3 plates were prepared for plasma and 5 for red cells. The most convenient method of replica plating was by using an Eppendorf Multipippette adapted to take disposable plastic tips. Microtitre plate sample positions and replicates were carefully indexed. Plates were sealed with sellotape before storing at -80°C. No observable change occurred in samples kept at this temperature for over 2.5 years.

2.2.2 Starch Gel Electrophoresis

The apparatus, buffer recipes and staining methods used were adapted slightly from those of other workers particularly Harris and Hopkinson(1978), Shaw and Prasad(1970) and Cole and Parkin (1981). They are briefly described below. Precise details of reagents are provided in Appendix 1.

Gel Preparation

All gels were made from a 12% w/v mixture of starch (Connaught) and gel buffer (as described in each recipe below). The starch was mixed well with the buffer in a flat-bottomed spherical flask, heated with constant mixing until becoming transparent shortly after coming to the boil, then degassed. As spherical flasks are liable to implode the flask was placed inside a safety tank during

degassing.

The hot, degassed starch solution was immediately poured into gel moulds consisting of a perspex former placed on a glass plate. The gel dimensions were 185 x 100 x 6mm. The starch solution is fairly viscous at this stage and was poured so that it formed a pronounced meniscus within the mould, slightly higher at one end. The top glass plate was then lowered onto the surface of the gel, starting from the higher end, taking care not to trap any air bubbles. The top plate was then pressed down, squeezing out any excess gel. Gels were usually prepared the day before use and left to set at room temperature. If gels were prepared the same day as required they were cooled to quicken setting.

Sample Treatment

In the case of samples being assayed for peptidase and esterase isozymes, it was found to be desirable to treat them prior to the electrophoretic run to counteract the effects of post-translational modification (see Section 3.1.3). The treatment consisted of the addition of 10µl of a 10mg/ml solution of dithiothreitol (DTT) to each 30µl sample, mixing, and incubation at 37°C for 1h.

Sample Application

Samples were applied to the gel using inserts cut from Whatman No. 3MM chromatography paper. The average size of inserts was about 4mm x 6mm, slightly less than the thickness of the gel. The gel was prepared by removing the top plate, drying the surface with tissue paper, running around the inside of the edge of the former with a scalpel blade to free the gel, and cutting 36-40 insert holes using a sharpened 3.5mm wide spatula. Inserts were blotted on

filter paper to remove excess liquid before being loaded into the gel.

Electrophoretic Apparatus

The apparatus was as illustrated in Whitehouse (1979). The loaded gel was supported horizontally on a copper coolant plate in turn supported on a glass plate spanning the buffer tanks. Current from a Heathkit IP-17 regulated d.c. power source connected to the tank electrodes was applied to the gel via folded J-cloth wicks. The surface of the gel was covered with a plastic sheet, taking care not to trap air bubbles, and a further coolant plate laid on top. Coolant at 4°C was circulated through the coolant plates. The whole apparatus was enclosed by a safety lid.

Running Conditions

The buffers, electrical conditions and running times for each of the 7 loci that were found to be polymorphic in blood (Chapter 3) are summarised in Tables 2.3 and 2.4. The voltages indicated are those supplied by the power-pack; actual voltages across the gel will be lower. The distance across the gel between the wicks was 75mm.

Staining Procedure

At the end of the electrophoretic run the gel was removed from the apparatus, the perspex former and inserts removed, and the gel surface blotted. A corner of the gel was cut to ensure correct subsequent orientation and the gel inverted onto a dry surface. The gel was then sliced equally using nylon fishing line stretched across spacers placed on each side, and surface moisture removed by

Table 2.3

Buffer systems used for main survey.

Code	Stock solutions	pH	Gel dilution	Bridge dilution	Ref.
1	0.5M NaH_2PO_4; 0.5M Na_2HPO_4:titrate	7.0	1:39	1:4	(1)*
2	0.245M NaH_2PO_4; 0.15M citric acid; +10M NaOH to raise pH.	7.0	1:79	1:0	(2)*
3	0.5M TRIS; 0.5M $NaOH_2PO_4$	7.4	1:99	1:4	(2)
4	0.5M TRIS; 0.07M citric acid; +6M HCl to reduce pH.	5.0	1:49	1:4	(3)
5	(i)0.06M LiOH; 0.229M boric acid. (ii)0.079M TRIS; 0.007M citric acid.	8.6	1Vol(i): 5.4Vol(ii)	(i)	(4)
6	0.5M TRIS +6M HCl to reduce pH.	8.0	(for staining solutions only)		

* Modified concentrations.

References

(1) Shaw & Prasad(1970).
(2) Harris & Hopkinson(1977).
(3) Cole & Parkin(1981).
(4) Gahne(1966).

Table 2.4

Running conditions for polymorphic systems used.

Enzyme/ Protein	E.C.	Buffer Code	Insert‡ position	V /V	I /mA	t /h	Migration
6PGD	1.1.1.44	1	N	110	100	5	Anodal
IDHC	1.1.1.42	2	N+1cm	110	75	4	Cathodal
PEPT, PEPD2, PEPD3	3.4.11 or 3.4.13.*	3	N	200	75	3.5	Anodal
EST2	3.1.1.1	4	N	200	75	3.5	Anodal
GP1	-	5	N	150 -300	75	4	Anodal

‡N = normal position, 2cm from cathodal wick.

blotting. Staining of the cut surface was carried out either by immersion in a solution of the reagents or by the application of a layer of stain contained in molten agar. Gels run for GPI (see Chapter 3) were stained using a filtered solution of 0.2% (w/v) Amido Black 10B and 0.4% (w/v) Nigrosine in a fixing solution of 50:50:10 methanol:water:acetic acid. A list of stain recipes for the other loci is provided in Table 2.5 (cf. Appendix 1).

Once the methodology for this project was finalised both gel slices were used only for TRIS-phosphate gels, where each half was stained for different peptidases. Gels run for 6PGD, IDH and the peptidases were incubated at 37°C until staining was sufficiently advanced to allow scoring (up to 3h). Esterases developed rapidly, and the reaction was stopped by washing in water as soon as the isozymes of interest were clearly visible (about 10 min). GPI was stained overnight then the background was destained by rinsing in water and then soaking in fixing solution for several hours, repeating this destaining procedure at least twice.

2.3 Data Handling and Statistical Analysis

The data were stored on a mainframe computer, and to minimise the possibility of errors each bird was identified by its unique BTO ring number in all data files. These identification numbers were annotated with tissue sample tray storage positions and, where appropriate, colour ring codes. Records were sorted according to sample tray locations, and hence gelling sequence, prior to the entry of electrophoretic data. The use of sorting routines to produce data listings in order of, for example, gelling sequence, colour ring codes or nest or clutch was found to be invaluable for minimising the occurrence of errors, and detecting any that arose.

Table 2.5

Staining Mixtures. (For GPI stain see text).

Reagent	Enzyme System				
	6PGD	IDHC	PEPD	PEPT	EST2
Buffer at r.t.	8ml 0.2M TRIS HCl pH8.0	6ml 0.5M TRIS HCl pH8.0	8ml 0.2M Phosphates pH7.0	8ml 0.2M Phosphates pH7.0	100ml 0.1M TRIS HCl pH7.0
Stock buffer (Table 3.2)	6	6	1	1	6 (N.B. lower pH)
Enzyme substrate	15mg 6-Phospho-gluconate	30mg Isocitric acid(Na_3)	20mg Leucyl-‡ tyrosine	10mg Leucyl-glycyl-glycine	40mg α-Naphthyl propionate in 2ml 50% acetone
Fast Blue RR Salt					25mg
Peroxidase* 5mg/ml			1.0ml	1.0ml	
Amino-acid oxidase* 4mg/ml			1.0ml	1.0ml	
3-Amino-9-ethyl carbazole** 8mg/ml			1.0ml	1.0ml	
0.2M $MgCl_2$	1.0ml	1.0ml			
NADP 4mg/ml	1.0ml	1.0ml			
MTT 10mg/ml	1.0ml	1.0ml			
PMS 2mg/ml	1.0ml	1.0ml			
2% Agar @ 70°C	12.0ml	12.0ml	12.0ml	12.0ml	
Total volume /gel	**24.0ml**	**22.0ml**	**23.0ml**	**23.0ml**	**102.0ml**

* in 0.2M Phosphates pH7.0 buffer.
** in acetone; added to staining mixture whilst stirring.
‡ stirred for 1h to dissolve.
See Appendix 1 for details of suppliers.

Keeping errors to a minimum is of special importance in family studies (Chapters 3 and 4). Separate data files were maintained for details of sightings, families and nest records. Nest records for each day were initially transferred to diskettes using a program written in BASIC for an Apple II microcomputer. The program was designed to minimise data preparation time. The nest data was subsequently transferred from diskette to the mainframe system. Programs were written in FORTRAN 77 to combine data from the different files, and to test pedigrees.

The statistics used are described at appropriate points in the text (see Section 5.2 for details of heterogeneity and goodness of fit tests, and the minimum expected values required in those tests). Many of the tables required, including those for observed genotype frequencies (Section 5.3.1) and estimation of disequilibria (Section 5.3.7), were constructed using the SPSS CROSSTABS program (Nie et al. 1975). Many statistical analyses were carried out using programs from SPSS (Nie et al. 1975) and BMDP (Dixon et al. 1983). Calculations such as exclusion probabilities (Chapter 4), Hardy Weinberg tests (Chapter 5) and genetic identity estimates (Chapter 6) were made using programs written in BASIC for the Apple II microcomputer. Statistical analyses involving small amounts of data were also made on the Apple II, normally using the STATSEASE statistical package written by Prof. B.C. Clarke (Genetics Department, University of Nottingham).

CHAPTER 3

THE ENZYME POLYMORPHISMS

3.1 Enzyme Systems

3.1.1 Previous Studies of House Sparrows

Starch gel electrophoresis has been used by a number of authors to investigate protein variability in the house sparrow and a summary of the tissue distributions of those loci found to be polymorphic is provided in Table 3.1. In calculating the proportion of polymorphic loci (P) most authors use Avise and Selander's (1972) level of 0.95, but neither the value of P nor its interpretation are being considered here. Consequently Harris and Hopkinson's (1972) definition of 'polymorphic' is used, referring to those loci where an allele exceeds a frequency of 0.99. Loci with more than one allele where the frequency of the commonest allele is greater than 0.99 are described as variable only.

Of the 13 polymorphic systems known to resolve well in liver tissue only 5 were present at sufficient concentration to be scored in blood. In addition, two plasma loci were found to be polymorphic and scorable. Details of the methods used during this study are provided in Chapter 2.

3.1.2 Description of Polymorphisms

The polymorphic enzyme loci investigated in this study will be described below. In all cases the fastest migrating allele has been designated A, the next fastest B, and so on. Consequently, in the case of 6PGD, IDHC and PEPD2 the allele designations differ from those of Parkin and Cole (1984a and pers. comm.), but agree for PEPT. Figures 3.1 - 3.12 contain photographs of the gels and

Table 3.1

Summary of polymorphic protein loci in the house sparrow as studied by various authors, indicating tissues used. All loci except where indicated were assayed by starch gel electrophoresis.

Locus	Study & Tissue[1]								
	Bush[2] et al. (1970)	Manwell & Baker (1975)	Klitz[3] (1972)	Fleischer[3] (1983)		Cole & Parkin (1981 & pers.comm.) & this study			
	S	**E**	**various**	**K**	**L**	**L**	**E**	**P**	**M**
PGM-1						G‡	N	N	N
PGM-2		4	M:G			G‡	N	N	G
ADA						G	P	N	G
6PGD		G	L:G			G	G*	N	
G6PDH		G[4]							
IDHA		G[5]			G[5]	G	P	N	
IDHC		G[5]			G[5]	G	G*	N	
ME		G				a	a		
EST1			(K:G[6])	G		G	N	N	G
EST2	a	a	La			a	N	G*	a
EST3	a		LKSa			a		a	
Fl-EST1					G	a			
PEPT						G	G*	G[7]	
PEPD2						G	G*	G[7]	G
PEPD3						G	G*	G[7]	G
SORDH						G‡	N	N	P
GP1	a	A:G	S:G[8]			N	N	G*	
ACON1[9]						G	N	N‡	P
Amylase[10]						a	N	a	

Key

G Good activity and resolution reported.
P Poor activity.
N No activity.
a Activity, but patterns poorly resolved or uninterpretable.
* Scored for all blood samples collected during this study.
‡ Assayed by isoelectricfocusing.

1 Tissues: S=serum P=plasma E=erythrocytes K=kidney L=liver M=muscle A=egg albumin.
2 cf. Bush (1964), Bush (1967).
3 cf. Klitz (1973), Johnston & Klitz (1977), Fleischer et al. (1983).
4 Refer to text.
5 IDHII, IDH2 = IDHA; IDHIII, IDH3 = IDHC; cf. text.
6 Designated 'Esterase II'.
7 Low activity but good resolution.
8 Designated transferrin; GP1 is probably transferrin (see text).
9 Sex-linked; cf. Baverstock et al.(1982).
10 Believed to consist of several loci.

diagrammatic interpretations of the zymograms.

Isocitrate Dehydrogenase (Cathodal Locus) (IDHC)

Under the conditions used, the dimeric product of this locus migrated cathodally (Figures 3.1 and 3.3). The locus has been investigated by a number of different authors who have referred to it using a variety of abbreviations: i.e. IDH II., IDH2, ICDH-M, ICDH-C and IDH-C (Table 3.1, Cole and Parkin 1984a). The locus will be referred to as 'IDHC' throughout this thesis. Activity was generally low, particularly when compared with that found in liver tissue, and some samples required incubation for up to 3h before achieving optimum staining. As activity varied among individuals and, to a lesser degree, among gels, many samples were assayed more than once to confirm scores. Activity differences among samples were more pronounced at this locus than at any of the other six investigated in detail. It was not established whether these differences were due to fluctuating efficiency of sample preparation, but some variation at least among samples would be expected to result from changes in erythrocyte concentration (Banerjee and Banerjee 1970) and the likelihood of cyclical variations in the levels of some enzymes (e.g. Brok-Simoni et al. 1976). In the more active $IDHC^{A}/^{A}$ homozygote types slight sub-banding occurred, including some at a similar position to the $IDHC^{B}$ allozyme, but no activity was observed in the position of the $IDHC^{A}/^{B}$ heterozygote's hybrid dimer, which was consequently regarded as diagnostic. $IDHC^{A}$ was the commonest allele.

Although staining was apparent, resolution at the more anodally-migrating locus, IDHA, was generally too poor to allow satisfactory scoring. IDHA produced the soluble form of the enzyme

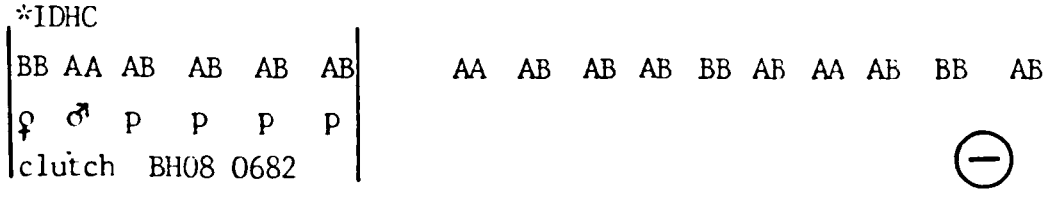


Figure 3.1 Photograph of IDH zymogram including one complete family, indicating IDHC genotypes. IDHC migrates cathodally from the insert line (i). p = progeny (nestling). cf. Figure 3.3.

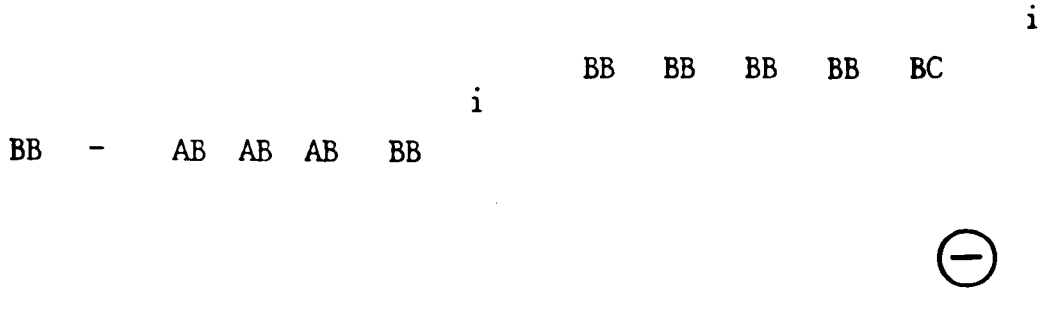
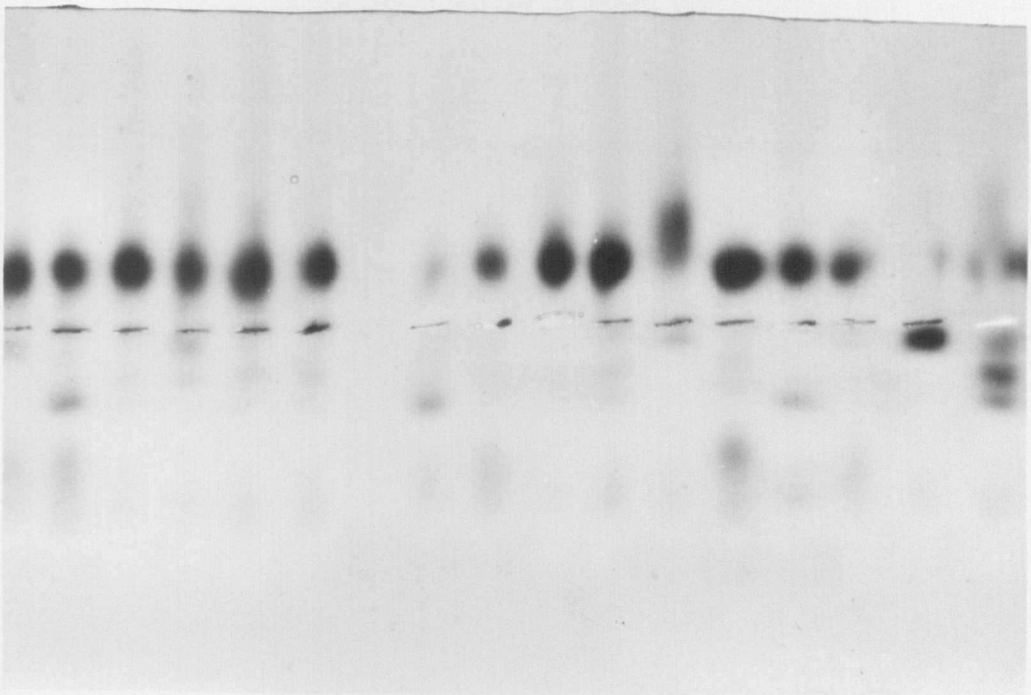
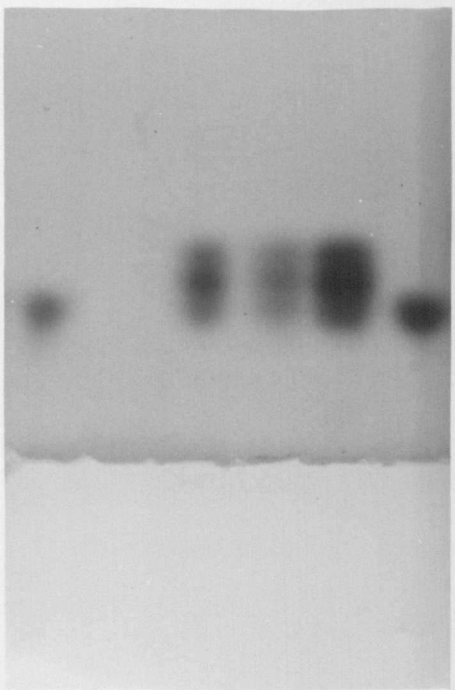
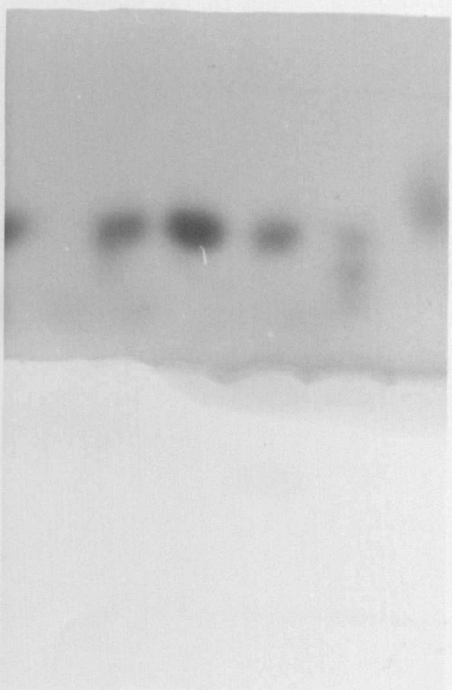


Figure 3.2 Photographs of 6PGD zymograms, showing genotypes. The gels have been cut along the insert line (i). cf. Figure 3.4.







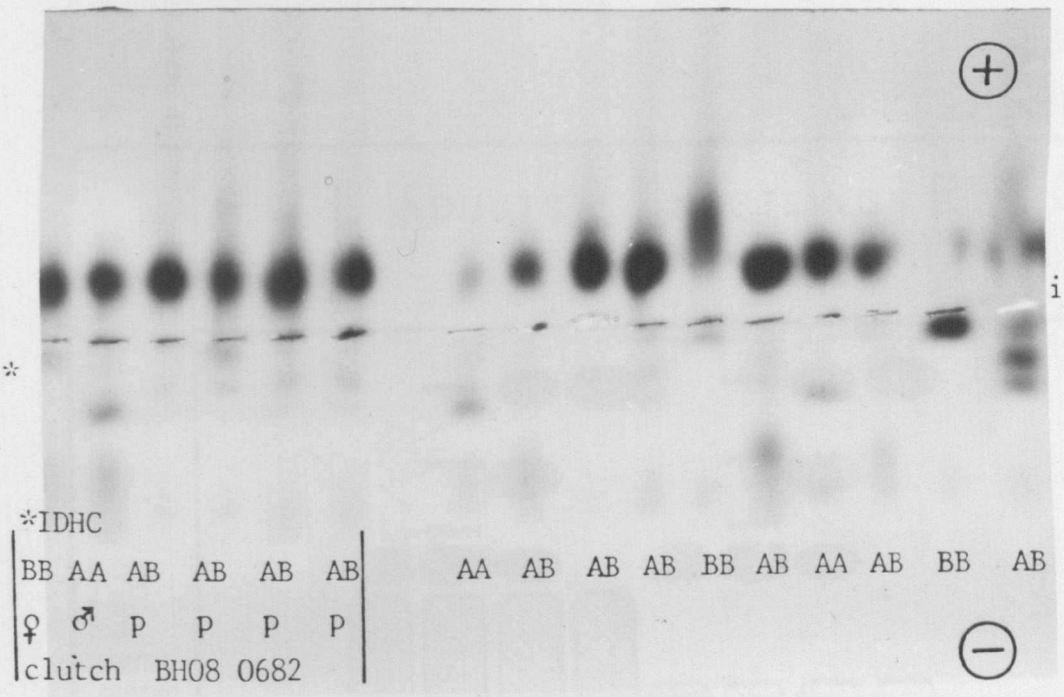


Figure 3.1 Photograph of IDH zymogram including one complete family, indicating IDHC genotypes. IDHC migrates cathodally from the insert line (i). p = progeny (nestling). cf. Figure 3.3.

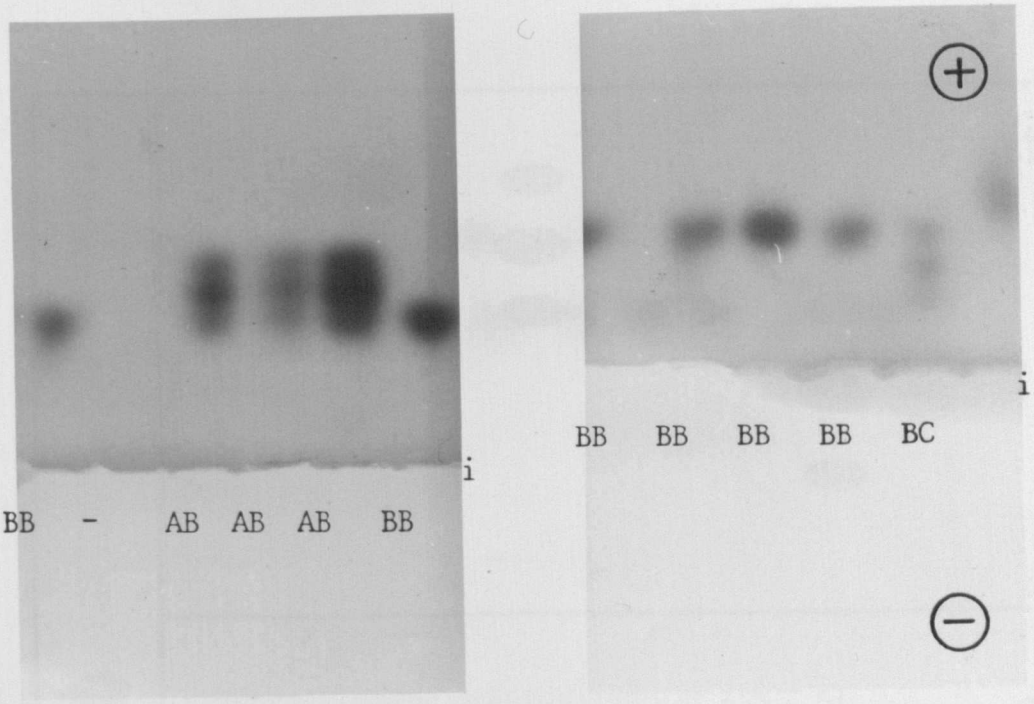


Figure 3.2 Photographs of 6PGD zymograms, showing genotypes. The gels have been cut along the insert line (i). cf. Figure 3.4.

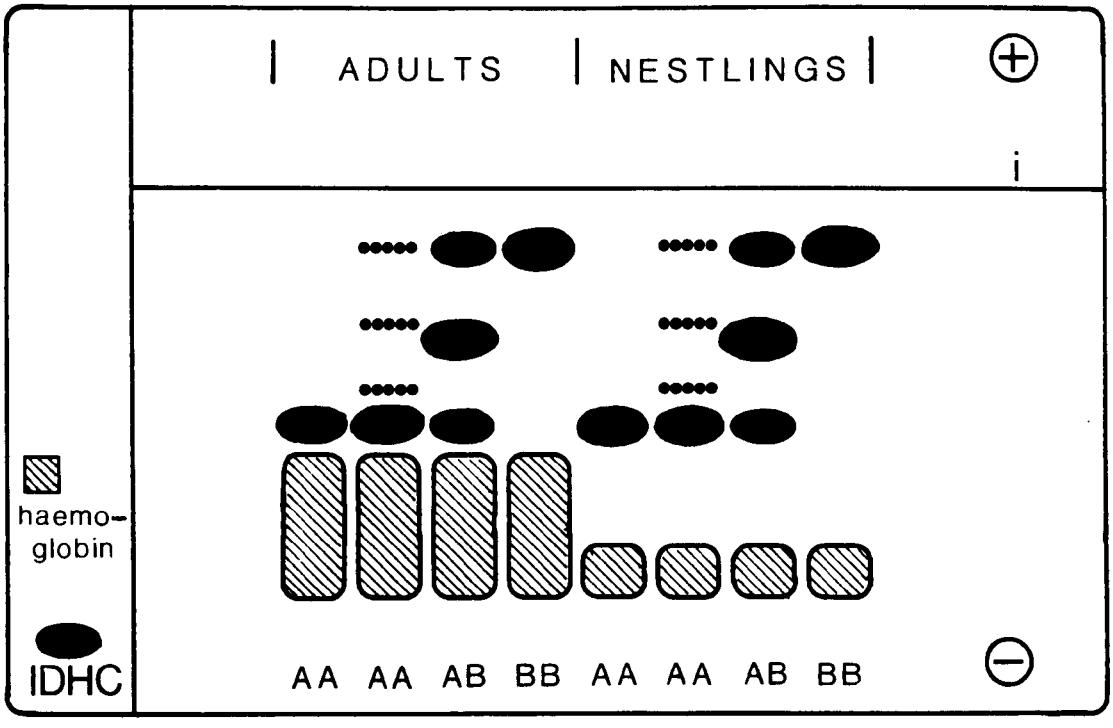


Figure 3.3 Diagram of IDHC zymogram indicating genotypes (Anodal IDH locus not shown). Differences between adult and nestling haemoglobin patterns are illustrated. i = insert line. cf. Figure 3.1.

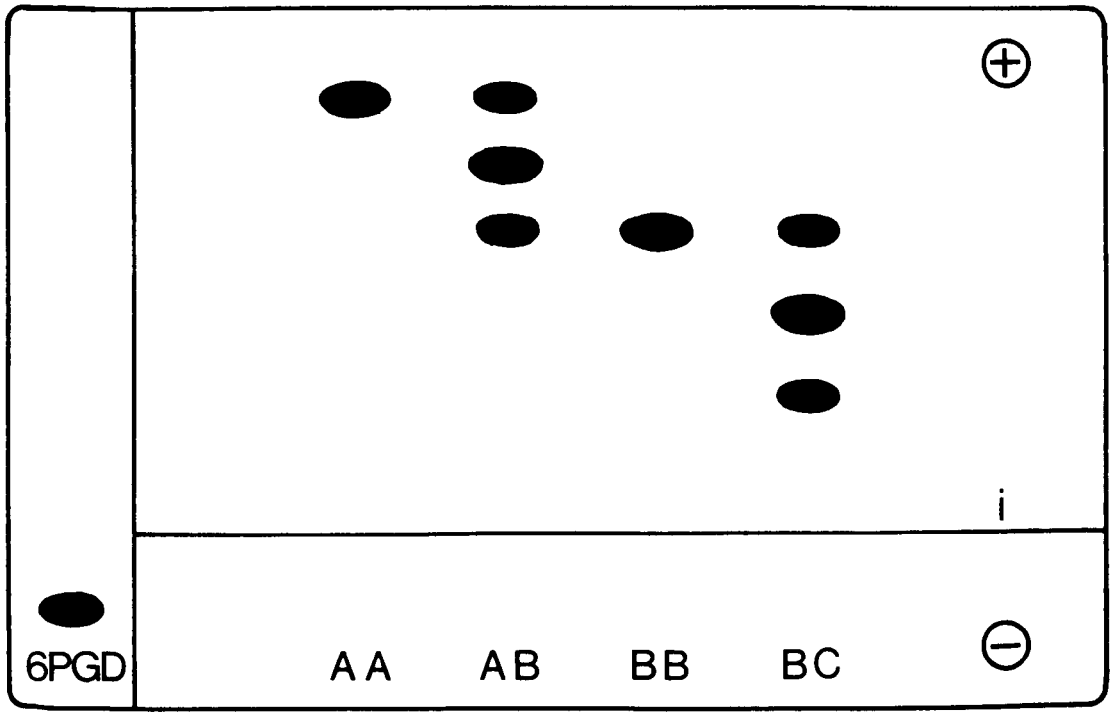


Figure 3.4 Diagram of 6PGD zymogram indicating genotypes. i = insert line. cf. Figure 3.2.

and IDHC the mitochondrial form (shown experimentally by S.R. Cole, pers. comm.).

6-Phosphogluconate Dehydrogenase (6PGD)

This locus is dimeric (three-banded heterozygotes); the $6PGD^{B}$ allele is the commonest (Figures 3.2 and 3.4). Shaw and Prasad (1971) recommend the addition of NADP to the gel and electrode buffers to obtain good activity, but this was found to be unnecessary.

Peptidases (PEPD3, PEPD2, PEPT)

The peptidases have been discussed in more detail by Cole and Parkin (1981). The peptidases were subject to a form of post-translational modification (see Harris and Hopkinson 1978) which resulted in smearing, mostly anodal to the normal band position, and poor resolution. This was found to be fully reversable by treatment with DTT (Section 2.2.2).

PEPT is monomeric (double banded heterozygotes) and the commonest allele was designated $PEPT^{D}$ (Figures 3.6 and 3.8). PEPD3 is monomeric and exhibits slightly lower activity than the dimeric PEPD2 locus. PEPD2 isozymes migrate to a position just cathodal of PEPD3 such that the $PEPD2^{A}/^{B}$ heterozygote overlaps the $PEPD3^{D}$ isozyme position (Figures 3.5 and 3.7). Consequently, only the faster migrating PEPD3 heterozygote, $PEPD3^{A}/^{B}$, could be scored with confidence in $PEPD2^{A}/^{B}$ types and so individuals having these PEPD2 genotypes were excluded from any further analyses concerning PEPD3. The common alleles were $PEPD2^{B}$ and $PEPD3^{B}$.

All the peptidase loci resolved well in plasma but activity was much lower than in erythrocyte extractions.

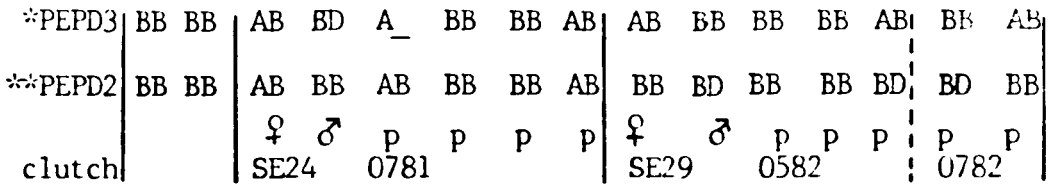


Figure 3.5 Photograph of dipeptidase zymogram including two families. Genotypes for PEPD3 and PEPD2 are indicated. A_ = presumed AD. p = progeny. Insert position cathodal, ie. to the bottom of the photograph. cf. Figure 3.7.

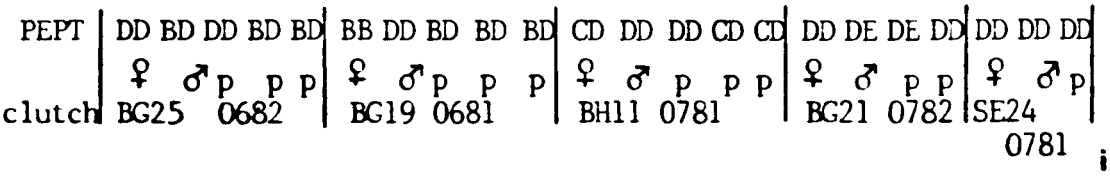
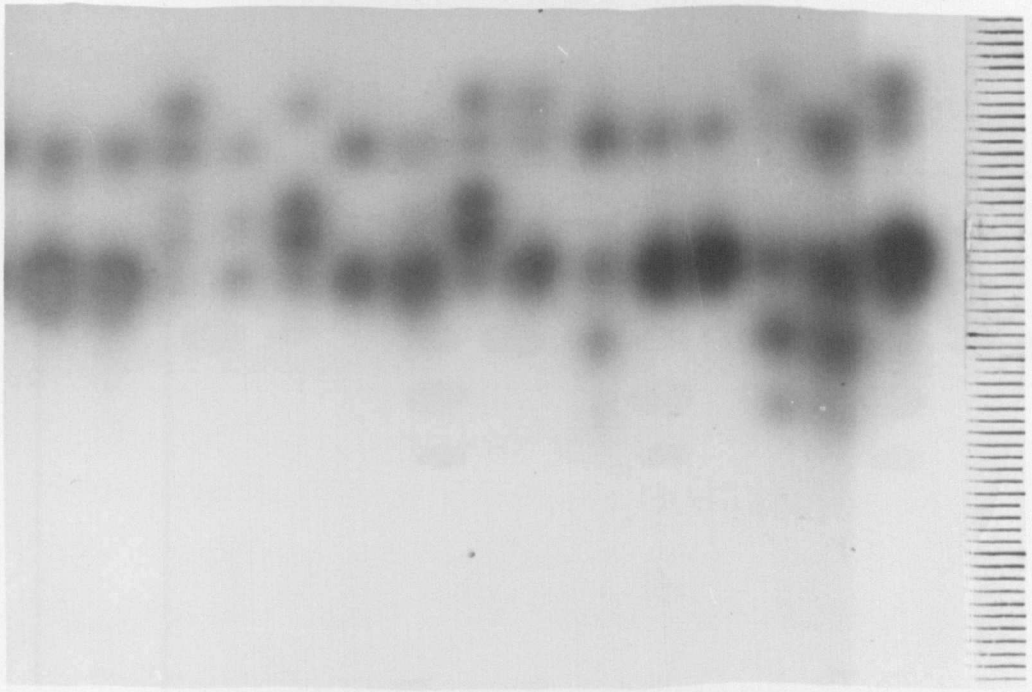
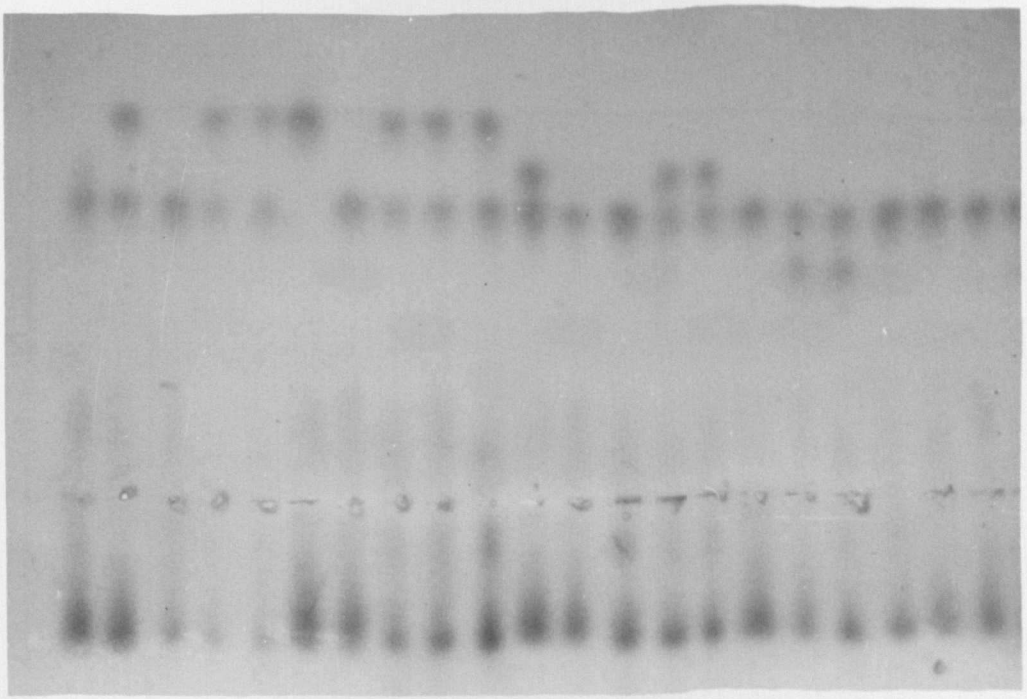


Figure 3.6 Photograph of PEPT zymogram, with genotypes for 5 families. p = progeny (nestling). i = insert line. cf.Figure 3.8.





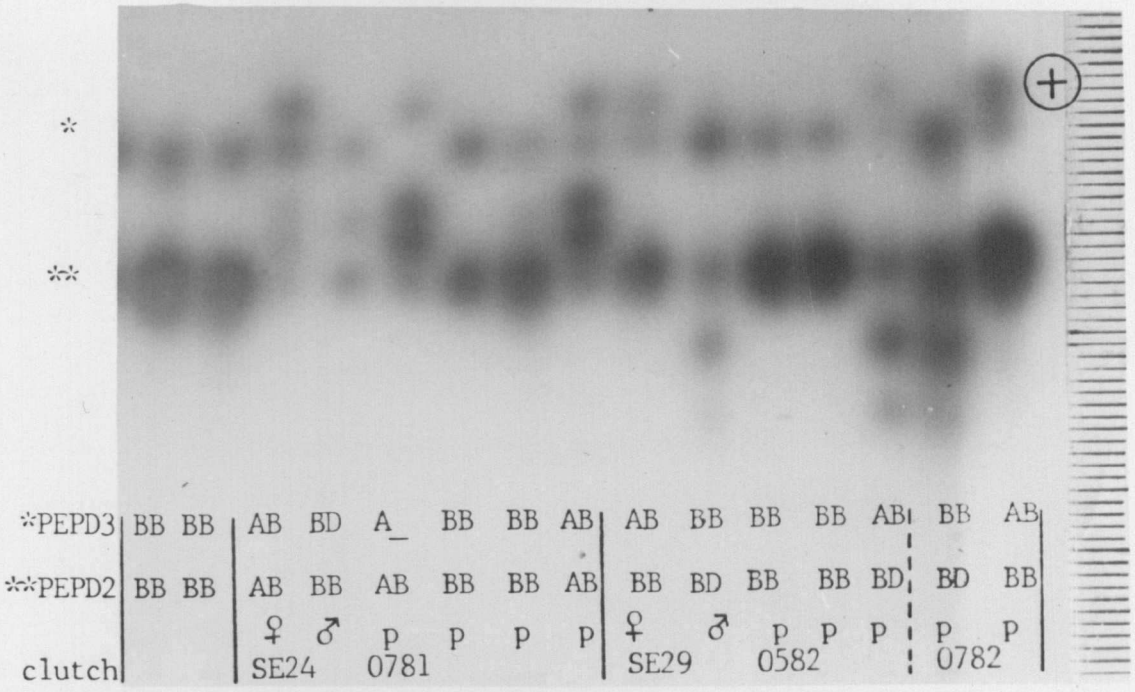


Figure 3.5 Photograph of dipeptidase zymogram including two families. Genotypes for PEPD3 and PEPD2 are indicated. A_ = presumed AD. p = progeny. Insert position cathodal, ie. to the bottom of the photograph. cf. Figure 3.7.

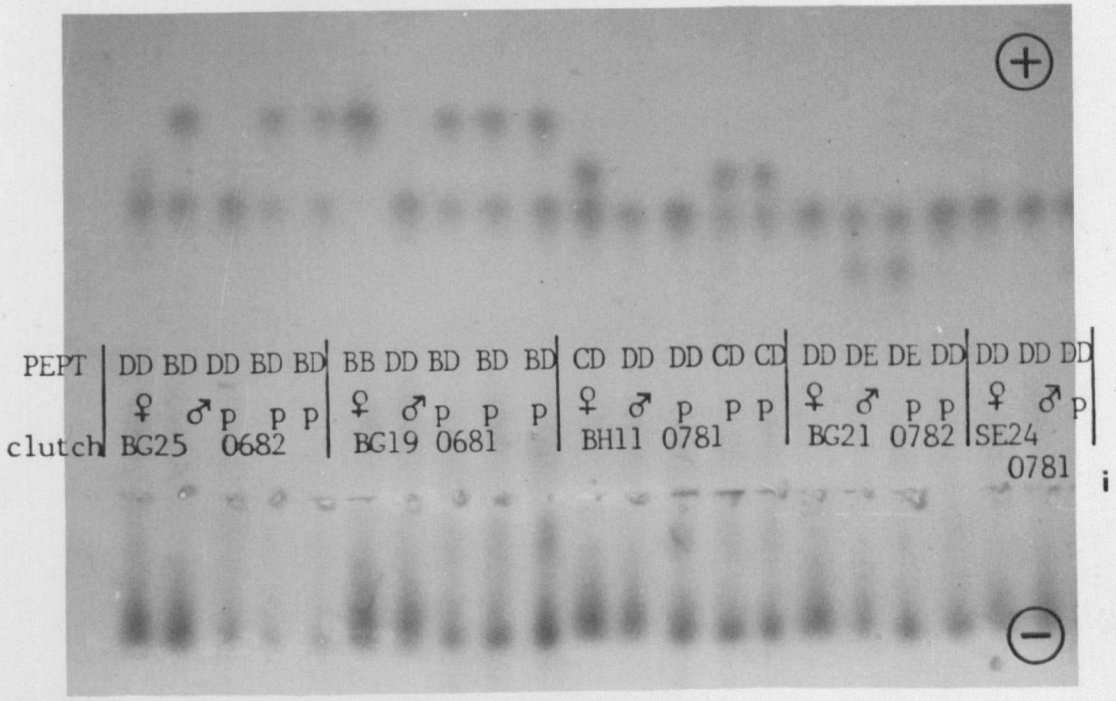


Figure 3.6 Photograph of PEPT zymogram, with genotypes for 5 families. p = progeny (nestling). i = insert line. cf.Figure 3.8.

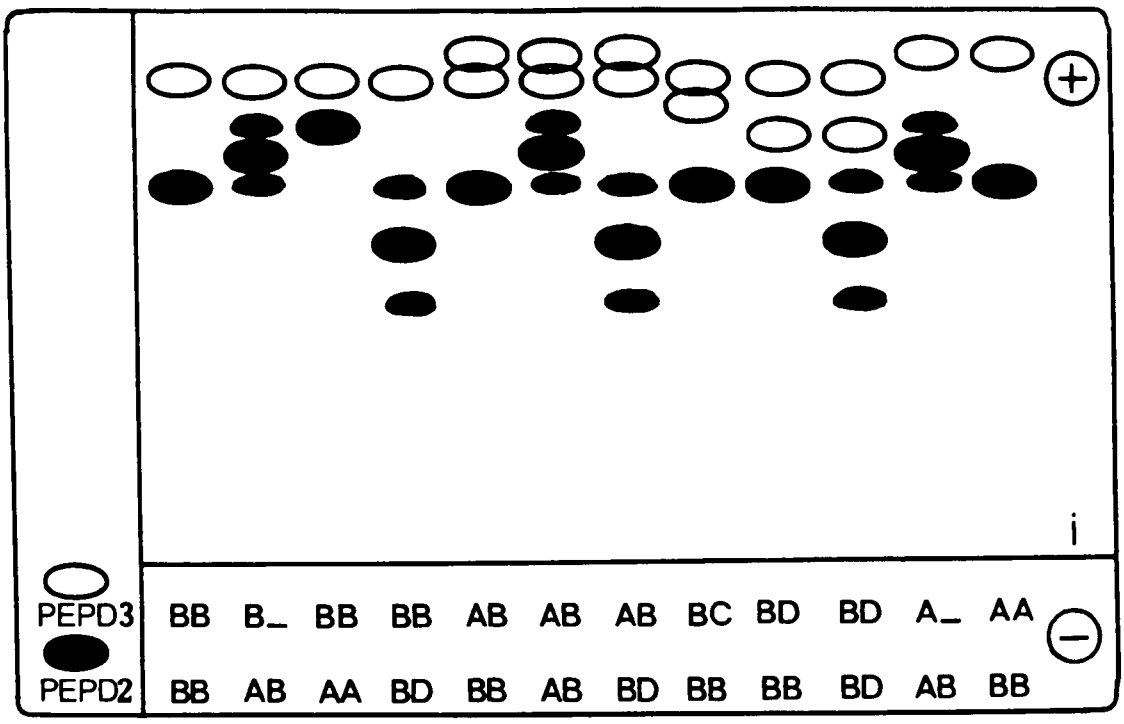


Figure 3.7 Diagram of dipeptidase zymograms indicating genotypes at each locus. i = insert line. cf. Figure 3.5.

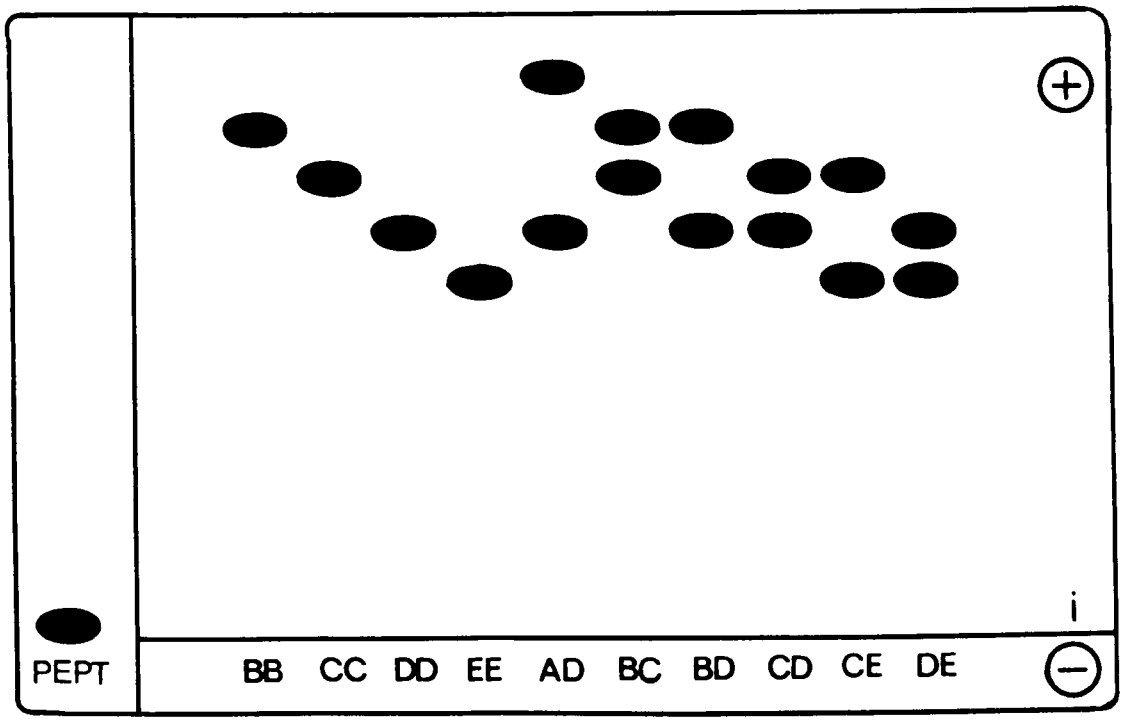


Figure 3.8 Diagram of PEPT zymogram indicating genotypes. i = insert line. cf. Figure 3.6.

General Protein 1 (GP1)

On the basis of comparison with published gel patterns depicting conalbumin bands (e.g. Sibley 1970 pp.13-15 and Figure 24; Baker 1965), published gel patterns for serum (or plasma) transferrin in various bird species (e.g. Montag and Dahlgren 1973), and the knowledge that, in birds, egg-white conalbumin is controlled by the same locus as serum transferrin (Ogden et al. 1962), the plasma locus described here as general protein 1 (GP1) is almost certainly transferrin. Polymorphism at this locus has been found for about 60% of bird species examined (Sibley et al. 1974); consequently most authors appear to rely on homologies of gel patterns with those for other species in the identification of transferrin.

Transferrin is an iron-binding protein and can therefore be detected by staining with a solution of Nitroso-R. This was used here, but was inconclusive as zones other than GP1 were also stained. Klitz (1972) presented data for a transferrin polymorphism in the house sparrow, but unfortunately included no details of the gel patterns or their interpretation. It is thought that Klitz's transferrin (as also studied by other workers at the same laboratory - refer to Table 3.1) is the same as GP1 here, but the genetic interpretations applied in each laboratory could conceivably be different. As it has not been conclusively proven that GP1 is transferrin the locus will be referred to throughout as GP1.

As has commonly been reported for transferrin (e.g. Montag and Dahlgren 1973, Ito et al. 1981), GP1 homozygotes show a 2-banded phenotype, and heterozygotes either 3- or 4-banded phenotypes. For

example, the faster-migrating band of the $\underline{GP1^{B}/^{B}}$ homozygote takes the same position as the slower band of the $\underline{GP1^{A}/^{A}}$ homozygote (see Figures 3.9 and 3.11). The faster-migrating component of each allele's gene products is generally regarded as being due to post-translational modification of the slower-migrating product, either by the addition of extra sialic acid residues (e.g. Ferguson 1980) or fewer Fe^{3+} ions (e.g. Stratil 1967). Addition of Fe^{3+} (Stratil op. cit.) in an attempt to saturate transferrin Fe-binding sites had no effect on mobility.

Of the two bands produced in the homozygote condition, the more anodally-migrating one was always less intense. Care had to be taken in scoring heterozygote types, therefore, as the most anodally-migrating band was then expected to be only half as intense as in the homozygote. In $\underline{GP1^{A}/^{B}}$ and $\underline{GP1^{B}/^{C}}$ heterozygotes the middle band was the most intense. A phenotype that at first sight appeared to have a more intense anodal band would generally, on further destaining or rerunning using a fresher sample, be revealed as such a heterozygote. This interpretation of patterns would appear to be sound as there were no more parent-offspring incompatibilities observed, relative to expectation, at this locus as compared with any other (Chapter 4). This locus was particularly sensitive to denaturation (resulting in reduced activity) following continued freeze-thawing or exposure to high temperatures. If any lysis of red blood cells occurred prior to their separation from the plasma component then the resulting extra proteins would obscure the banding patterns for GPl, preventing it from being scored. Extreme care was therefore taken to avoid lysis and to minimise the number of freeze-thaw events (Chapter 2).

No other interpretable variation was observed in gels stained

Albumen

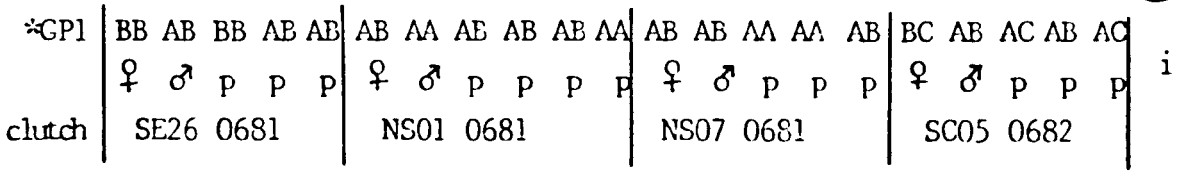


Figure 3.9 Photograph of gel stained for general proteins. The GP1 genotypes of 4 complete families are shown. p = progeny (nestling). i = insert line. cf. Figure 3.11.

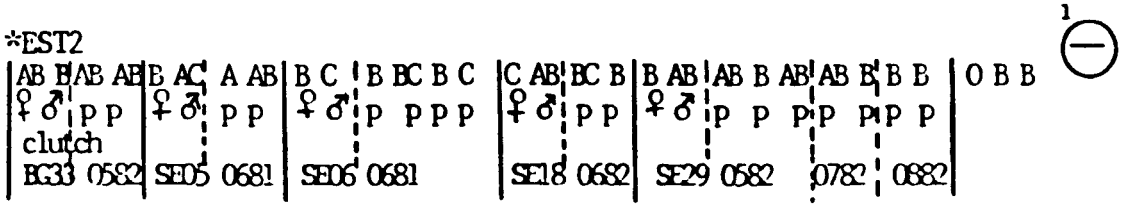
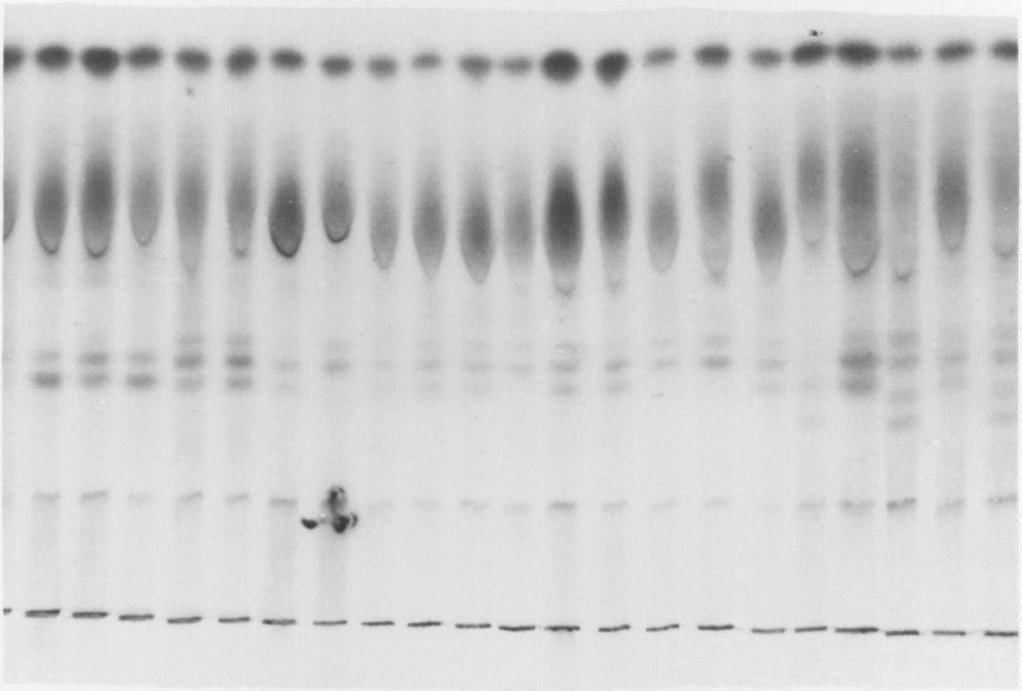
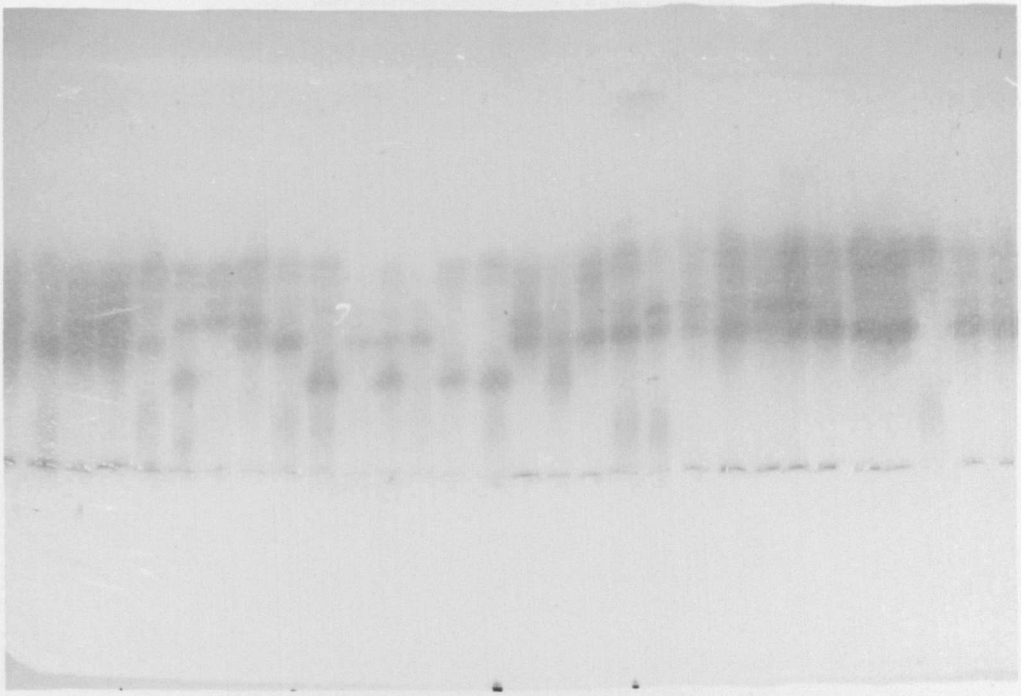


Figure 3.10 Photograph of esterase zymogram including 5 complete families. The EST2 phenotypes are shown. p = progeny (nestling). i = insert line. cf. Figure 3.12.





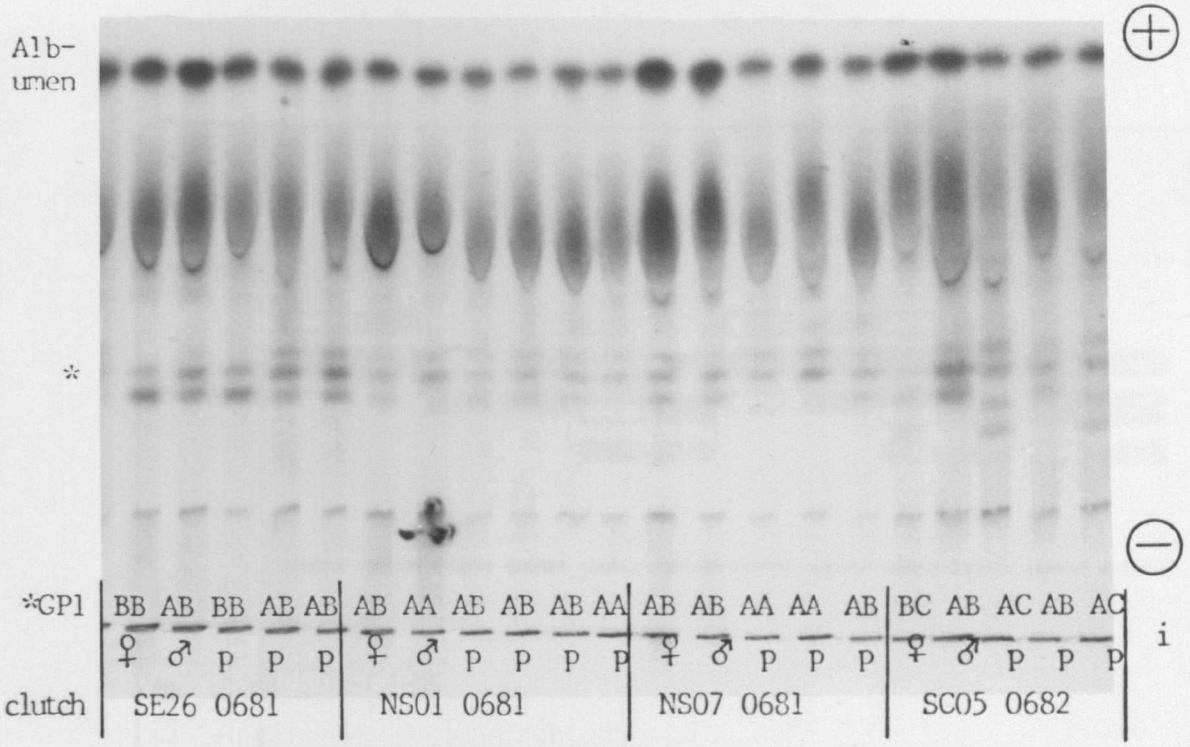


Figure 3.9 Photograph of gel stained for general proteins. The GP1 genotypes of 4 complete families are shown. p = progeny (nestling). i = insert line. cf. Figure 3.11.

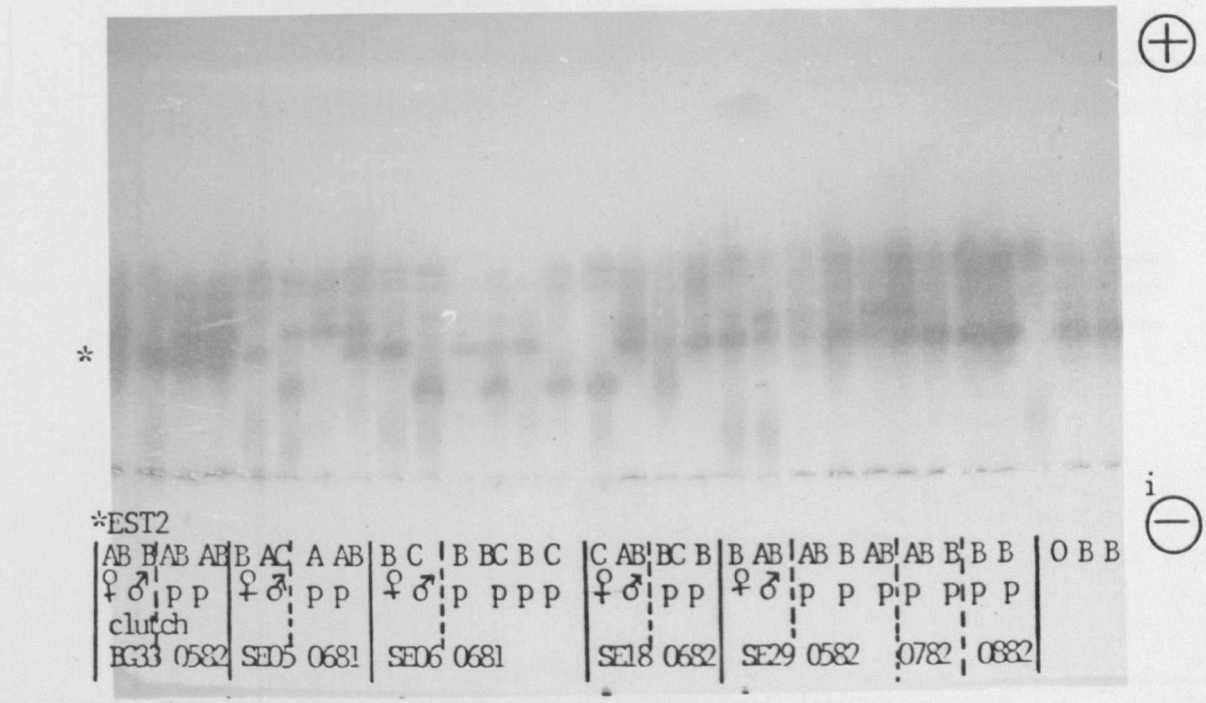


Figure 3.10 Photograph of esterase zymogram including 5 complete families. The EST2 phenotypes are shown. p = progeny (nestling). i = insert line. cf. Figure 3.12.

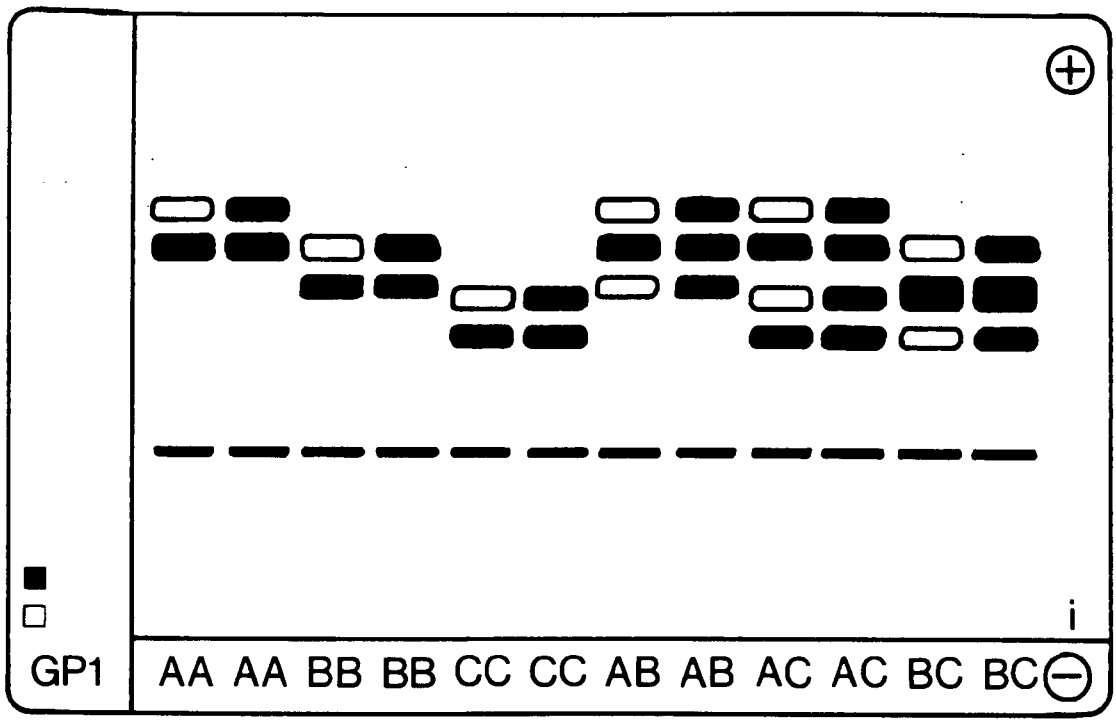


Figure 3.11 Diagram of general protein bands indicating GP1 genotypes. Proteins banding anodal to GP1 are not illustrated. i = insert line. cf. Figure 3.9.

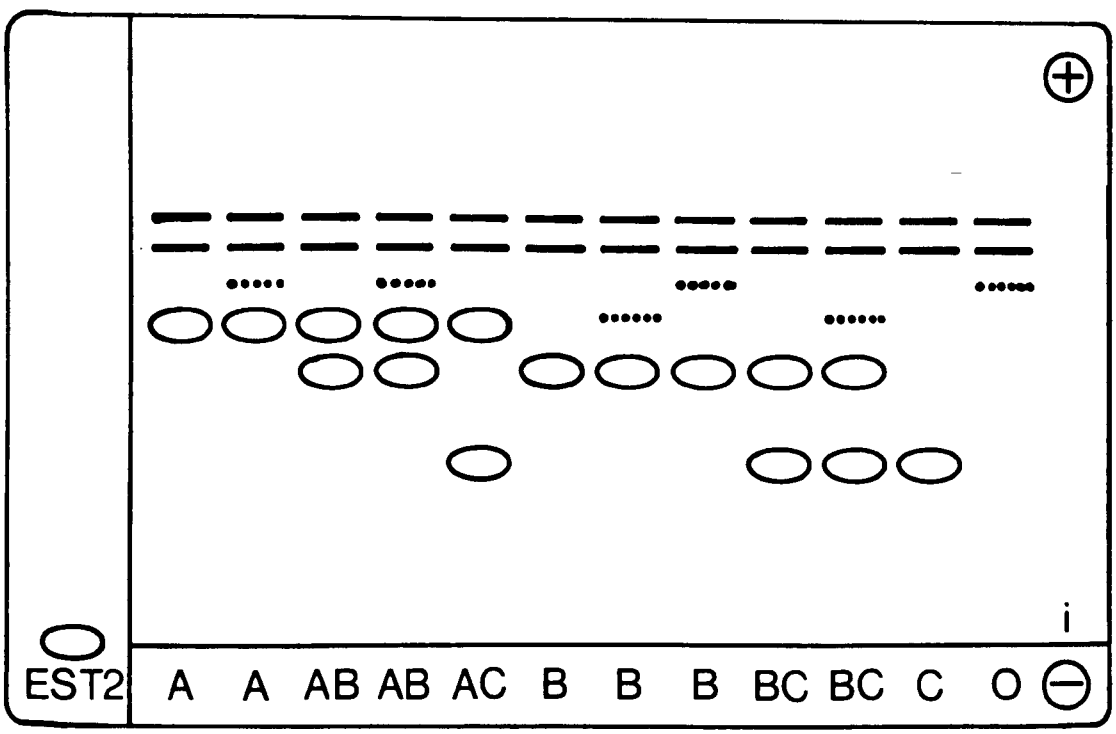


Figure 3.12 Diagram of esterase zymogram, indicating EST2 phenotypes. An esterase locus (or loci) migrating anodally to EST2 is also illustrated. i = insert line. cf. Figure 3.10.

for general proteins, except for one probable albumin heterozygote among the 1500 or so individuals examined.

Esterase (EST2)

EST1, which is polymorphic in liver (Table 3.1), is not present in plasma. The slower of the anodally-migrating esterases, the previously incompletely resolved product of the locus designated EST2 and suggested to be monomorphic (Cole and Parkin 1981), was found to be polymorphic in plasma. A number of substrates were used to investigate the plasma esterase loci. Essentially the same relative activities of the various bands persisted, and activity was strongest with α-naphthyl propionate. The polymorphic locus of interest was immediately cathodal to a less rapidly staining region which apparently consists of two monomorphic, but poorly resolved, bands. Resolution of EST2 was improved slightly by treatment with DTT (see Section 2.2.2).

The commonest allele at this locus was $\underline{EST2^B}$. Null, or silent, alleles are those either producing no gene product or a product of such low activity that it is not detectable by electrophoretic techniques. Evidence for the presence of a null allele (designated here as $\underline{EST2^0}$) was provided by individuals which exhibited no activity at this locus, breeding data (see below) and an apparent pronounced deficiency of heterozygotes (excess of the rarer homozygote types) when Hardy-Weinberg expectations were calculated without allowance for the existence of nulls (see Section 5.3.2). Individuals of genotype $\underline{EST2^{0/0}}$ showed normal activity and patterns for the slower staining, faster migrating, esterase locus (or loci). They also showed apparently normal concentrations for other proteins found in plasma, as indicated by staining intensity

on general protein gels.

3.1.3 Other Loci Investigated

The relative paucity of loci detectable in blood as compared with other tissues has already been mentioned. All loci for which good electrophoretic systems had already been found in other tissues (Table 3.1) were investigated. Of the known polymorphic loci, Esterase-1 (EST1), sorbitol dehydrogenase (SORDH) and the more anodally-migrating aconitase locus (ACON1) showed no activity in blood tissues, and adenosine deaminase (ADA) showed activity in the erythrocyte preparations of some individuals only. Phosphoglucomutase (PGM) was investigated using both starch gels and isoelectricfocussing (LKB 1977). Only one locus was apparent in erythrocytes (PGM3), and one variant was detected amongst 20 individual haemolysates; this is probably the same locus for which Manwell and Baker (1975) reported two variant types in 25 individuals. Amongst additional loci, glucose-6-phosphate dehydrogenase (G6PDH: using the same conditions as for 6PGD) appeared to be invariant in the erythrocyte fraction for 323 individuals.

It is not known why the variability reported by Manwell and Baker (1975) for G6PDH, with allele frequencies of 0.42/0.58, was not detected here. The activity loss reported by Manwell and Baker for stored tissue samples was not observed, even following storage for at least one year at -80^{o}C. Interestingly, an analysis of Manwell and Baker's data indicates that the genotypic frequencies presented for G6PDH in their sample departed from Hardy-Weinberg expectations, with an excess of heterozygotes ($\underline{G}$ = 4.68, $\underline{p}$<0.05). No similar departure has been recorded for any other locus (Cole and

Parkin 1981).

Amylase in plasma was investigated using isoelectricfocusing and extremely variable patterns (14-21 bands per individual) were apparent but unfortunately could not be interpreted satisfactorily.

In addition to the 18 monomophic or slightly variable loci detected in erythrocyte preparations and listed by Cole and Parkin (1981) should be added glutathione reductase (GSR) and NADPH diaphorase (DIA-NADPH). These were visualised simultaneously on the spare slice from gels run for IDH. Both loci showed slight variation only. During the initial consideration of the suitability of different tissues for sampling (see Chapter 2) femoral and pectoral muscle tissues were tested for activity of the polymorphic loci known in other tissues. The use of isoelectricfocussing for SORDH (Cole and Parkin 1981) revealed only slight activity in some individuals and none in the remainder, whilst the same individuals showed good activity in liver samples. As GPI and EST2 could not be assayed in muscle preparations it was concluded that muscle would have no net advantage over blood as a source of suitable proteins.

3.2 Segregation Analyses

Before using the genotypic classes presumed to code for the observed isozymes for any further genetic analyses it was desirable to check that their inheritance agreed with a Mendelian model and that the loci were independent. For this purpose the data collected from 126 clutches (42 at Brackenhurst, 82 at Sutton Bonington and 2 at Nottingham) where both parents and offspring were sampled has been used.

3.2.1 Mendelian Ratios

The total observed allozyme types of the offspring produced by each mating class for each locus are given in Tables 3.2 and 3.3. Note that a few individuals have occasionally not been scored for a particular locus because of, for example, insufficient material being collected or, very occasionally, damage to a particular sample. The maximum number of offspring omitted is 6 (in the case of GPl: due to lysis of red cells), except for PEPD3 where several clutches have been omitted because of possible interference by PEPD2 heterozygotes (see above).

The occurrence of null alleles for EST2 has been discussed (above). Expected segregation ratios have not been estimated for most mating types at this locus as it was not possible, given the sample sizes within each category, to estimate the frequency of unidentifiable null heterozygotes with a useful level of confidence. The informative crosses concerning the allelic nature of the uncommon $\underline{EST2^{C}}$ are presumed to have each involved an $\underline{EST2^{C/O}}$ parent (see Table 3.3); this genotype is expected to be much more common than the homozygote for $\underline{EST2^{C}}$.

The table includes any incompatible offspring types that were detected. Within any mating class, with the possible exception of the $\underline{IDHC^{A/B}}$ category, their numbers are very small relative to the compatible offspring. In many instances where one scored parental genotype was in some respect incorrect the assessment of segregation ratios, providing that alleles at the locus segregate in a Mendelian fashion, will not be substantially affected. For the $\underline{IDHC^{A/B}}$ class the 10 such offspring were contained in just 4 clutches. In the case of $\underline{IDHC^{A/A}}$ x $\underline{IDHC^{B/B}}$ matings an incorrect assessment of parentage is more likely than usual to result in the appearance of

Table 3.2

Summary of total progeny genotypes observed for each class of mating at those loci where each phenotype was interpreted genotypically. The deviations from expected ratios (shown in parentheses) among compatible offspring are indicated. n.s. = not significant.

Locus	Mating type	Number of clutches	Total offspring genotypes	Deviation
6PGD			AB BB	
	AB X BB	8	12 11	n.s.
			(1 : 1)	
	BB X BB	118	1* 340	-
			(1)	
PEPD3‡			AB BB BC BD DD	
	AB X BB	6	8 7 - 1* -	n.s.
			(1 : 1)	
	AB X BD	1	1† 1 - 1 -	n.s.
			(1 : 1 : 1 : 1)	
	BB X BB	71	- 207 - 5* -	-
			(1)	
	BB X BC	2	- 4 2 - -	n.s.
			(1 : 1)	
	BB X BD	25	- 51 - 20 1*	G = 14.0
			(1 : - : 1)	$p < 0.001$

* incompatible types

‡ excludes several instances where heterozygote for PEPD2 might interfere with scoring of PEPD3.

† excludes 1 'AZ' where PEPD2 type was AB and therefore obscured PEPD3 position.

Table 3.2 cont'd

Locus	Mating type	Number of clutches	Total offspring genotypes				Deviation
PEPD2			AA	AB	BB	BD	
	AB X AB	1	-	-	1	-	n.s.
			(1 : 2 : 1)				
	AB X BB	21	-	32	22	-	n.s.
				(1 : 1)			
	BB X BB	103	-	3*	298	-	-
					(1)		
	BB X BD	2	-	-	4	1	n.s.
					(1 : 1)		
IDHC			AA	AB	BB	AC	
	AA X AA	36	103	3*	1*	-	-
			(1)				
	AA X AB	49	69	68	2*	-	n.s.
			(1 : 1)				
	AA X BB	13	6*	33	4*	-	-
				(1)			
	AB X AB	19	16	25	8	-	n.s.
			(1 : 2 : 1)				
	AB X BB	8	1*	12	10	-	n.s.
				(1 : 1)			
	BB X BB	1	-	-	2	-	-
					(1)		

* incompatible types

Table 3.2 cont'd

Locus	Mating type	Number of clutches	Total offspring genotypes							Deviation
PEPT			AD	BD	CC	CD	DD	DE	BC	
	BB X DD	2	-	6 (1)	-	-	-	-	-	-
	BD X DD	13	-	15 (1 :	- :	- :	22 1)	-	-	n.s.
	CD X DD	10	-	-	1*	14 (1 :	9 1)	-	-	n.s.
	BC X CD	1	-	3 (1 :	1 :	1 :	- :	-:	1)	n.s.
	DD X DD	97	-	1*	-	3*	281 (1)	-	-	-
	DD X DE	3					6 (1 :	2 1)		n.s.

*incompatible types

Table 3.2 cont'd

Locus	Mating type	Number of clutches	Total offspring genotypes						Deviation
			AA	AB	AC	BB	BC	CC	
GP1									
	AA X AB	2	2	4	-	-	-	-	n.s.
			(1 :	1)					
	AA X AC	1	3	-	1	-	-	-	n.s.
			(1 :	- :	1)				
	AA X BB	2	1*	6	-	-	-	-	-
				(1)					
	AB X AB	9	3	12	-	12	-	-	n.s.
			(1 :	2 :	- :	1)			
	AB X BB	41	2*	51	1*	63	2*	-	n.s.
				(1 :	- :	1)			
	AB X BC	7	-	8	3	7	4	-	n.s.
				(1 :	1 :	1 :	1)		
	AC X BB	5	-	10	-	-	2	-	G = 5.81
				(1 :	- :	- :	1)		$p < 0.025$
	BB X BB	53	-	4*	-	140	1*	-	-
						(1)			
	BC X BC	1	-	-	-	1	2	-	n.s.
						(1 :	2 :	1)	
	BB X BC	5	-	-	-	12	3	-	G = 5.78
						(1 :	1)		$p < 0.025$

* incompatible types

Table 3.3

Summary of total progeny phenotypes for each class of mating at the locus known to have null alleles (EST2). The classes A, B, C will include both the appropriate heterozygous null and homozygous genotypes. Expected ratios (in parentheses) could be inferred in two instances only. n.s. = not significant.

Locus	Mating type	Number of clutches	Total offspring phenotypes							Deviation
			A	AB	B	BC	C	AC	O	
EST2										
	AB X AB	3	2	2	2	-	-	-	-	n.s.
			(1 : 2 : 1)							
	AB X B	27	1	38	45	-	-	-	-	-
	B X B	88	-	5*	246	-	-	-	1	-
	AB X C	1	1	-	1	1	-	1	-	-
	AB X O	1	-	-	2	-	-	-	-	n.s.
			(1 : - : 1)							
	AC X B	1	1	2	-	-	-	-	-	-
	B X C	1	-	-	2	1	1	-	-	-
	A X B	2	-	1	1	-	-	-	-	-
	B X O	2	-	-	7	-	-	-	-	-

* incompatible types.

parent-offspring incompatibilities. No particular conclusion need necessarily be drawn, then, concerning any differences in inheritance for this mating. The proportion of incompatible offspring, at 4.6%, is higher at IDHC than any of the other loci. This result is also expected to be a function of genotype frequency. These considerations, together with a discussion concerning the possible origin of incompatibilities, will be found in detail in Chapter 4.

With regard to the compatible offspring genotypes significant deviations from expected Mendelian ratios were found for one PEPD3 mating type ($p < 0.001$) and two GP1 mating types ($p < 0.025$: refer to Table 3.2). Ideally, every possible mating type would be examined at a locus in a large number of cases, but in practice many mating types will be relatively rare. In order to maximise the information obtained concerning segregation ratios the data were further sorted according to the frequency of transmission of segregating parental (heterozygote) genotypes to the progeny (Table 3.4). Each parental allele is expected to be transmitted at an equal rate. Matings where neither parent is heterozygous will be uninformative, whilst one of each pair of offspring alleles will be informative if one parent is apparently heterozygous and both alleles if both parents are heterozygous. Possible null heterozygotes at the EST2 locus were ignored.

Significant deviations from a 1:1 ratio were again found for PEPD3 and GP1 only, indicating a deficiency of $\underline{PEPD3}^{D}$ ($p < 0.001$), an excess of $\underline{PEPD3}^{B}$ ($p < 0.001$), a deficiency of $\underline{GP1}^{C}$ ($p < 0.001$) and an excess of $\underline{GP1}^{B}$ ($p < 0.01$). Reference to Table 3.2 reveals that the highly significant deficiency for $\underline{GP1}^{C}$ is contributed towards by all parental matings that include the allele. The

Table 3.4

Summary of observed segregations among offspring of heterozygous parents, for each allele considered separately. (Goodness of fit test refers to an expected ratio of 1:1. n.s. = not significant.)

Locus	Heterozygous parental allele	Allele observed in offspring	Alternative observed in offspring	Goodness of fit
6PGD	A	12	11	n.s.
	B	11	12	n.s.
PEPD3	A	9	9	n.s.
	B	62	30	G = 11.4 ***
	C	2	4	n.s.
	D	21	53	G = 14.3 ***
PEPD2	A	32	24	n.s.
	B	28	33	n.s.
	D	1	5	n.s.
IDHC	A	138	119	n.s.
	B	121	97	n.s.
PEPT	B	18	22	n.s.
	C	14	15	n.s.
	D	40	31	n.s.
	E	2	6	n.s.
GP1	A	91	116	n.s.
	B	145	104	G = 6.78 **
	C	15	44	G = 14.89 ***
EST2	A	50	55	n.s.
	B	55	47	n.s.
	C	0	3	n.s.

** $p < 0.01$ *** $p < 0.001$

excess of transmitted $GP1^B$ types would appear to be associated with a deficiency of $GP1^A$ alleles (though not significant) as well as a deficiency of $GP1^C$. As discussed above, GP1 is almost certainly transferrin (Section 3.1.3). Another study of birds (ring-necked pheasants Phasianus colchicus) has also reported a distorted segregation ratio at the transferrin locus (Vohs and Carr 1969). A higher mortality of individuals carrying the Tf^A allele was suggested, but no possible reason for the maintenance of this allele was given. A similar situation was found regarding an esterase locus in starlings (Sturnus vulgaris), where the commonest allele was transmitted significantly more often than expected in backcross matings (Evans 1980). This was interpreted as being the result of selection during the pre-fledging period favouring the commonest homozygote type.

There are numerous potential explanations to account for the results concerning house sparrows, ranging from methodological factors to actual fitness differences. I shall first consider the possibility that individuals having the deficient alleles were not always detected on gels.

With regard to the allele exhibiting a deficiency at each locus, the observed rate of transmission is in each case close to half of expectation. If the deficiencies were due to mis-scoring of gels, then the implication would be that, as on average only half are detected, as many parents would have possessed these alleles as were actually scored as having them (there was no indication that protein concentrations/activities were any different for adults and nestlings at these loci). On this basis one would expect some 26 $PEPD3^D$ and 20 $GP1^C$ alleles to go undetected among the parents. One would in turn expect as many of their offspring to be scored for

the allele as there were alleles detected among the offspring of detected parents. These would in almost all instances be incompatible with the recorded mating type. This is patently not the case (Table 3.2); a few incompatibilities are accounted for by these alleles but this is not unexpected when the overall observed incompatibility rates are considered (see Chapter 4).

Polymorphic loci which modify certain proteins in an electrophoretically detectable way are known from studies of Drosophila (Cochrane and Richmond 1979). Providing that the sampling of clutches (and therefore parents) is random, the frequency of a phenotype produced by the combined effects of a modifier locus and the visualised protein locus is expected to be the same in both the entire parental and entire offspring groups. This will apply regardless of whether the action of the appropriate alleles at the modifier locus is dominant or recessive. Thus, if the phenotype was subject to genetic modification then many more incompatibilities would have been expected than were observed. Their relative absence strongly supports the conclusion that the results are not accounted for by a polymorphic modifier locus. Similar considerations apply also to non-genetic sources of modification, and a similar conclusion reached.

More elaborate hypotheses to explain the departure from expected ratios could, of course, be constructed. For example, it might be suggested that the penetrance of the alleles increases with age, so that all appropriate adults are detected but not all nestlings. In the small number of instances where nestlings or juveniles were retrapped and retested as adults there was no difference observed in the isozymes, but the numbers are too few to be conclusive. There is no evidence from other studies to suggest

that changes might occur with age, except for ontogenic ones early in the development of nestlings where the relative expressions at different loci coding for a particular class of protein may change (e.g. haemoglobin: Kostelecka-Myrcha *et al.* 1971; LDH: Bush 1967). These latter are believed to be related to the changing physiological requirements of a bird at fledging and are not known to occur with respect to alleles at an individual locus.

The possibility of some kind of selective difference affecting alleles at PEPD3 and GP1 must therefore be considered. There are many distinct opportunities for the operation of selection prior to the sampling of chicks. A simple comparison of mean fledging success per clutch (from data in Table 3.2) between parents heterozygous for the alleles showing deficiency and those homozygous for the alleles with segregation excess revealed no significant differences. Observed fledging success was very slightly higher for parents possessing the $GP1^{C}$ allele as compared with the remainder, and there was no measurable difference between $PEPD3^{B/D}$ and $PEPD3^{B/B}$ parents. Thus the distorted segregation ratios could not be explained by a simple model of fitness differences among eggs or nestlings unless this were compensated for by increased clutch size or substitution for zygotes of low fitness prior to egg-laying. There was no evidence for a difference in clutch-size. Eggs in this species are each ovulated immediately following the laying of the previous one, and each yolk is believed to take 4 days prior to ovulation to develop (Schifferli 1976). Departures from a laying schedule of one egg per day are very rare (Seel 1968a). It is unlikely therefore that any deaths of zygotes prior to egg laying are compensated for by the laying of additional eggs.

The lack of any difference in fledging success need not be

incompatible with an hypothesis of post-zygotic selection acting in a density-dependent fashion. Mortality in house sparrow clutches is known to have a density-dependent component (see Section 1.4). The modal number of eggs laid per clutch during this study was 4, agreeing with observation elsewhere in England (e.g. Summer-Smith 1963). With a mean fledging number close to 3, mortality in successful clutches approaches 25%. Thus post-zygotic selection could explain the distorted segregation ratios if the selection were operating in a threshold manner upon the deficient genotypes. If this were the case, then those broods where no mortality was observed prior to sampling would not be subject to the density-dependent selection, and no deviant segregation ratios would be expected in those broods. There were 35 clutches without mortality where both parents were sampled. Segregation distortion was found to apply to these broods to a similar extent to that already observed in the entire sample (Table 3.5).

There was no evidence, then, for simple density-dependent selection, but the possibility of frequency-dependent selection has not been excluded. It might be hypothesised that selection against genotypes including the $\underline{PEPD^{B}}$ and $\underline{GPI^{C}}$ alleles was stronger in broods containing more of those genotypes. This hypothesis would be compatible with the observation of distorted segregation ratios in broods not exhibiting mortality; these broods could simply be those where no selection would have occurred because of the low initial frequency of the alleles. Clearly, the distorted ratios demonstrate that if frequency-dependent selection were operating then it was not balanced. Further, no evidence was obtained from the nestling period to suggest that the affected alleles were ever at a selective advantage. If frequency-dependent selection, or some combination of

Table 3.5

Segregation of alleles in parents with genotypes $\underline{PEPD3}^{\underline{B/D}}$ or $\underline{GP1}^{\underline{C/R}}$ (where R refers to A or B), for clutches exhibiting no mortality.

Locus	Allele		Goodness of fit
	B	D	
PEPD3	27	11	$G_1 = 6.95$, $p < 0.01$
	C	R	
EST2	7	26	$G_1 = 11.64$, $p < 0.001$

frequency- and density-dependent selection were operating, then the strength of this selection might be expected to depend upon environmental conditions, and to vary both spatially and temporally. There was, however, no heterogeneity for the segregations among the site or year classes for either locus (Table 3.6).

Pre-zygotic, or gametic, selection can also lead to distorted ratios. It has only occasionally been observed, and may be particularly difficult to distinguish from other forms of selection (see Cavalli-Sforza and Bodmer 1971). In most examples of pre-zygotic selection, including the t complex in Mus (see Bennett 1975), the SD locus in Drosophila melanogaster (see Hartl 1977) and the D factor in Aedes aegypti (Hickey and Craig 1966), segregation distortion occurs in one sex only. There was no significant heterogeneity between the sexes, however, for either of the loci showing distortion in house sparrows (Table 3.6).

Whatever form of selection was operating, it was clearly unbalanced prior to fledging in the cases of both loci under discussion. The maintenance of these polymorphisms implies that some other factor was at work at a later stage in the life cycle. Some possibilities are investigated and discussed elsewhere (Chapter 5).

3.2.2 Linkage Tests

With respect to the possibility of sex-linkage, the presence of several heterozygote types in the heterogametic female sex at all the loci indicated that none was sex-linked. The house sparrow is reported as having 38 pairs of chromosomes (Castroviejo et al. 1969, Bulatova et al. 1972) and close linkage between any of the 7 loci under study was,therefore, not expected. The tests that were

Table 3.6

Segregation at PEPD3 and GP1 loci compared across sites, sexes and years. For GP1, R refers to $\underline{GP1^A}$ and $\underline{GP1^B}$. Comparisons were made using heterogeneity G tests with Williams' correction (Sokal and Rohlf 1981).

Locus	Comparison		Allele transmitted		
PEPD3			B	D	
	Sites †	SB	25	13	
		BR	25	5	$G_1 = 2.66$
	Sexes	Female	27	8	
		Male	24	12	$G_1 = 0.94$
	Years	1980	8	3	
		1981	23	4	
		1982	20	13	$G_2 = 4.42$
GP1			C	R	
	Sites	SB	8	30	
		BR	7	14	$G_1 = 1.01$
	Sexes	Female	10	32	
		Male	5	12	$G_1 = 0.19$
	Years	1980	2	1	
		1981	8	29	
		1982	5	14	$G_2 = 2.23$

† excludes one informative clutch at Nottingham site.

carried out for the presence of linkage will be described in this section.

There were several instances where the same pair were known to fledge more than one clutch successfully and, therefore, in testing for linkage between loci the data concerning successive offspring, which would be full sibs, were combined. Clutches containing incompatibilities (see above) were excluded from this analysis. Even though in most instances the remainder of the offspring in the clutch will be correct (Chapter 4) it was considered safer to omit them as a small number of errors can be very misleading (in wrongly excluding close linkage). The remaining clutches may have contained undetected non-relatives or other errors at a low frequency (Chapter 4).

As there are likely to be differences in crossover frequency (Cavalli-Sforza and Bodmer 1971, p.872) linkage should ideally be analysed separately for each sex, but as the amount of data was small no attempt was made to do this here. In linkage analyses only a proportion of matings are informative. Specifically, for codominant systems such as the enzyme loci being examined, informative families will be those containing at least one parent which is heterozygous for both of the loci being investigated and, as the phases of the parents are unknown, have at least two offspring. In order to determine the phase (in coupling or in repulsion), the genotypes of a double heterozygote parent's parents would need to be known; this information was never available in this study. Where, as here, the family sizes are small, conclusive data concerning even close linkage will rarely be obtained from individual families and the probabilities of the observed segregations in several families must be combined.

Morton (1955) has shown that the most efficient method of detecting linkage from a succession of family samples is by the use of a sequential probability ratio test (Wald 1947). The procedure for the application of the test, together with tables of the log probability ratio (lod) scores which it uses, have been fully described (Morton 1955, 1957; Maynard Smith, Penrose and Smith 1961). The tables have been constructed with respect to man where the prior probability of randomly selected loci being linked is taken as 0.05. The house sparrow has many more chromosomes than man (see above), but the total includes many very small ones. The use of the formulae derived using the prior probability for humans is therefore regarded as reasonable for an initial analysis.

In the case of enzyme loci showing complete penetrance and no dominance, the estimation of lod scores is particularly easy. As all families included in a test are selected on the basis of the genotypes of the parents ('complete selection') and there is no dominance there will be no bias in the analysis and consequently no corrections to the lods will be required. Even where there are null alleles, as in the case of EST2 here, there will be no bias provided that the selection of double heterozygotes does not include parents known to be heterozygous for a null allele (as deduced from the genotypes of their offspring). In this way, if a selected family includes an intercross where the non-selected parent is heterozygous for a null allele then scoring will automatically take place as it would for a backcross where segregation is analysed only in the selected parent; the situation is analagous to that for ABO blood groups in man (Maynard Smith, Penrose and Smith, 1961, p.52).

Only matings that include one parent which is doubly heterozygous for the loci being tested are informative. Most

informative house sparrow matings were of the double backcross or backcross-intercross type. Heterozygous offspring for 'two-way' intercross matings (e.g. AB from ABxAB - refer to Table 4.3) are not informative. At least two informative offspring per family are required for the test. There were a few double intercross matings (both parents double heterozygotes); both parents may provide linkage information in such crosses (see Morton 1957: p.56). The lod score ($\underline{z}$) appropriate to the absolute numbers of possible 'coupling' and 'repulsion' events for the value ($\underline{\theta}_1$) of the linkage coefficient ($\underline{\theta}$) being tested ($\underline{H}_1$: $\underline{\theta} = \underline{\theta}_1$) was obtained from the tables (e.g. from Morton 1955: Table 10). Lod scores for one large family (11 informative offspring), and, where required, for very small values of $\underline{\theta}_1$, were found using the formulae provided by Maynard Smith *et al.* (1961).

Conventionally, if the sum of sequentially encountered lod scores, $\underline{Z}(\theta_1) \geq 3$ then the null hypothesis of no linkage ($\underline{H}_0$: $\underline{\theta} = 0.5$) is rejected. If $\underline{Z}(\theta_1) \leq -2$ then the actual recombination fraction, $\underline{\theta}$, will be significantly greater than the tested value $\underline{\theta}_1$.

If the loci are linked, the best estimate of $\underline{\theta}$, $\hat{\underline{\theta}}$, is obtained where the value of $\underline{Z}(\hat{\theta})$ is maximised.

The lod scores obtained are given in Table 3.7. From the observation of segregations within families, extremely close linkage ($\underline{\theta} \simeq 0$) could be immediately excluded from most combinations of loci. In the case of four pairs of loci there were no informative matings. Only one informative mating was available concerning the possibility of linkage between the two dipeptidase loci, PEPD2 and PEPD3, because the common heterozygote for PEPD2 interferes with scoring at PEPD3 (Section 3.1.2). For several pairs of the more variable loci the possibility of linkage was excluded for values of

Table 3.7

Lod scores Z(θ) for sequential linkage tests.

Locus pair	Informative families	Informative offspring	θ						
			0.00	0.01	0.02	0.03	0.04	0.05	0.10	0.20
6PGD-PEPD3	0									
6PGD-PEPD2	1	3	$-\infty$	-1.402						
6PGD-IDHC	4	12	$-\infty$	< -2	< -2	< -2	< -2	< -2	-1.339	
6PGD-GP1	0									
6PGD-PEPT	0									
6PGD-EST2	1	2 *								
PEPD3-PEPD2	1	4 *								
PEPD3-IDHC	3	14	$-\infty$	< -2	< -2	< -2	< -2	< -2	-1.7	
PEPD3-GP1	5	25	$-\infty$	< -2	< -2	< -2	< -2	< -2	< -2	-0.369
PEPD3-PEPT	1	3	$-\infty$	-1.402						
PEPD3-EST2	2	6	$-\infty$	< -2	< -2	< -2	< -2	-1.906		
PEPD2-IDHC	6	18	$-\infty$	< -2	< -2	< -2	< -2	< -2	-1.983	
PEPD2-GP1	0									
PEPD2-PEPT	2	5 *								
PEPD2-EST2	4	19	$-\infty$	< -2	< -2	< -2	< -2	< -2	-1.028	
IDHC-GP1	15	51	$-\infty$	< -2	< -2	< -2	< -2	< -2	< -2	-1.554
IDHC-PEPT	4	11	$-\infty$	< -2	< -2	< -2	< -2	< -2	-1.561	
IDHC-EST2	6	25	$-\infty$	< -2	< -2	< -2	< -2	< -2	< -2	-1.810
PEPT-GP1	4	9	$-\infty$	< -2	< -2	< -2	-1.946			
PEPT-EST2	4	10	$-\infty$	-0.229						
GP1-EST2	4	13	$-\infty$	-1.339						

* No observed 'recombination' events, but data very limited: $0 < Z(\theta) < 1$ for all values of θ.

$\underline{\theta}$ of up to 0.10. As the tested value for $\underline{\theta}$ is increased, disproportionately more information is required to reject the $\underline{H}_1$ hypothesis. It was not possible to exclude the possibility of close linkage for three pairs of loci with informative matings.

The possibility of selection, which may be gametic, affecting certain alleles at the PEPD3 and GP1 loci has been discussed above (Section 3.2.1). The similarity of the effects upon segregation ratios at each of these loci might be explained if the loci were in linkage disequilibrium, possibly both with some other unknown locus, and consequently indirectly subjected to the same selection. Linkage disequilibrium is more likely to occur when loci are linked (perhaps in a chromosomal inversion), and whether these loci are linked is therefore of particular interest. It was concluded, however, that these loci were not closely linked (Table 3.7: $\hat{\underline{\theta}} > 0.10$). The possibility of linkage disequilibria in these study populations will be dealt with elsewhere (Chapter 5).

3.3 Summary

Techniques were established which allowed seven polymorphic protein loci to be sampled non-destructively in two natural populations of house sparrows. A detailed investigation of the genetics of these loci was conducted for 126 complete families containing 363 offspring.

There were a number of instances where offspring phenotypes were judged to be genetically incompatible with their parents, but there was general agreement with the model of inheritance proposed for each locus. Segregations at four loci (6PGD, PEPD2, PEPD3 and IDHC) agreed with a simple Mendelian model of codominant

inheritance. One locus (EST2) was interpreted as having three codominant alleles and one or more recessive null alleles. Two loci (PEPD3 and GP1) showed segregation distortion in all sex, site and year classes. This distortion was not attributable to the misinterpretation of gel patterns; possible causes involving the operation of natural selection were discussed.

Linkage analyses allowed the possibility of close linkage to be eliminated for 14 out of 21 pairs of loci, including the pair showing distorted segregation ratios. Few informative matings were observed for the remaining 7 locus pairs. No significant evidence was obtained for linkage between any combination of loci.

CHAPTER 4

THE BREEDING SYSTEM

4.1 Introduction

This chapter is concerned with behavioural and genetic aspects of the breeding system of the house sparrow particularly as inferred through the analysis of the genotypes of families by electrophoresis. It is intended to also include an assessment of the problems and the general applicability of the technique. In particular, the number of parents responsible for producing a clutch is discussed; consideration of the randomness or otherwise of mating is made elsewhere.

A variety of studies of mating systems that have been made using genetic markers have been reviewed in Section 1.3. In particular, two studies of house sparrows were referred to. In one of these (Manwell and Baker 1975), evidence was obtained to suggest that intraspecific brood parasitism had occurred in the clutches of two different females. The other (Cheke 1969) provided evidence for the occurrence of successful extrabond copulation between a male tree sparrow Passer montanus and a female tree sparrow which was (unusually) mated to a male house sparrow. Mate guarding has been described in many animals, particularly bird species (see Hoogland and Sherman 1978 for examples), and the report of an apparently successful extrabond copulation might suggest that mate surveillance is limited in house sparrows.

Attempted extrabond copulations have been recorded in a variety of avian species which otherwise appear to be monogamous, particularly in the larger and more readily observable, colonial ones (see Gladstone 1979 and Wickler and Seibt 1983 for reviews).

Extrabond copulations are frequently referred to in terms of 'promiscuity' (e.g. Gladstone 1979, Stacey 1982), but 'promiscuity' may be misleading as it more exactly refers to those mating systems where mating occurs spontaneously and at random (Wittenberger 1979). In a majority, at least, of these examples, attempted extrabond copulations appear to be forced by males upon reluctant females (e.g. Coombs 1960, Meanley 1955, Fujioka and Yamagishi 1981, Kushlan 1973, MacRoberts 1973, Werschkul 1982). Attempts by males to fertilise extra females in addition to their mates are frequently interpreted as an adaptive consequence of a lower requirement for investment by males, as compared with females, in their individual offspring (e.g. Beecher and Beecher 1979, Trivers 1972). It is not known whether these attempts are successful in many species, but they are known to achieve success in captive mallards (Burns et al. 1980). Successful extrapair copulations have also been recorded in lesser snow geese (Mineau and Cooke 1979), and, following the vasectomy of resident territorial males, in red-winged blackbirds (Bray et al. 1975). This latter observation is perhaps the most surprising as many workers have marked individual red-winged blackbirds but never observed extrabond copulations, and only rarely seen males enter the territories of other males. The result illustrates the difficulties of observing such behaviour directly in the wild, and the possibility of its unexpectedly widespread occurrence.

In view particularly of the reports concerning house sparrows referred to above, use was made of seven polymorphic enzyme loci to investigate the possible occurrence of extrapair fertilisations or intraspecific brood parasitism in a large numbers of complete families in two study populations. Offspring phenotypes judged to

be genetically incompatible with those of their parents are referred to here as 'mismatches' (following Ashton 1980). The detection of a mismatch in a family implies that one or both putative parents have been 'excluded' from the possibility of parentage; the detection is an 'exclusion' event. Exclusions may be made in either of two ways:

(i) - an offspring possesses an allele absent from both putative parents,

(ii) - an offspring lacks an allele it is expected to possess from knowledge of the putative parents' genotypes.

Additionally, exclusions among putative siblings (usually due to multiple paternity) may be deduced from the collective genotypes of the sibship. Where dominance occurs, as for example in several blood group systems, sibship examination might provide additional exclusions in instances where offspring do otherwise, when considered individually, appear to be compatible with their parents; no extra exclusions will be obtained in this way when codominant enzyme systems are used if both putative parents are tested.

Mismatches may occur as a consequence of either sample identification or phenotype determination errors, or else as a result of any of a number of behavioural events. For example, a rare null allele in a parent might lead to an exclusion on the basis of criterion (ii). It is with regard to possibilities of this sort that European forensic experts insist on the use of criterion (i) alone in human paternity cases (Salmon, Seger and Salmon 1980). Behavioural events which might produce a mismatch include extrapair copulations (or more generally 'extrabond' copulations, allowing for

behaviourally identifiable polygamous breeding systems), inadvertent or deliberate adoption of another's offspring, and mate changing. The probability of detecting nonparentage, as demonstrated by a mismatch, resulting from such behavioural events is described as the exclusion probability.

4.1.1 Exclusion Probabilities

The probability of exclusion (P_i) for each marker locus depends on the number of alleles and their frequencies at that locus. In general, a locus is more useful the more alleles it has, and maximum efficiency is theoretically achieved when the frequencies of each allele at a locus are the same (Selvin 1980). The combined exclusion probability P_E for n unlinked systems is given by the formula:

$$P_E = 1 - \prod_{i=1}^{i=n} (1-P_i)$$

The exact calculation of an exclusion probability depends on the precise nature of the events of which detection is possible. For example, if the parentage of both parents is being tested then the exclusion probability is calculated as the probability of detection through mismatches of a randomly selected adult male-female pair. If only paternity is in doubt, then only the probability of exclusion for a randomly selected male adult is considered. The exclusion of paternity or else of parentage are the two probabilities of most interest; where the gene frequencies are similar in each sex, the probability for exclusion of a false mother is expected to be the same as for a false father. In general, nonpaternity and nonparentage are the possibilities

considered as occurrence of wrong maternity but correct paternity, in the putative pedigrees studied, is thought unlikely. The probability of exclusion of a false parent will increase with the number of affected offspring tested.

The calculation of an exclusion probability requires the assumption that all genotypes are phenotypically identifiable. The formulae in general use make the further assumption that the populations from which families are to be tested are in Hardy-Weinberg equilibrium. Jamieson (1965) provided a formula for the probability of exclusion of paternity, using both exclusion criteria described above, by systems with two or more codominant alleles. This was extended to the general multiple offspring case by Gundel and Reetz (1981), who made the further assumption that the entire litter would have been sired by the same male. Gundel and Reetz also provided a formula for the probability of exclusion of parentage in the multiple allele, single offspring case, an extension of Nielsen's (1970) two-allele method. These formulae are applied below in an attempt to investigate the frequency of nonpaternity/ nonparentage in house sparrow clutches.

Foltz (1981a) has provided a maximum-likelihood method for the estimation of nonpaternity, but this will not be applied here as the required assumptions that all putative mothers are true mothers and that all offspring in a clutch are sired by a single male may not be fulfilled. Foltz also assumed that no errors were made in the identification of individuals or in the handling and typing of tissue samples. Most workers recognise that such errors will occur, and Ashton (1983) reduced his exclusion estimates accordingly. Lathrop et al. (1983) recently presented a maximum-likelihood model for the estimation of laboratory error but made the

assumption of no 'field' error (i.e. in the attribution of offspring to parents). As it is probably unreasonable to make such an assumption in an ecological study of the kind described here, the formal maximum-likelihood approach will not be applied; a consideration is made, however, of possible sources of error and their consequences.

4.2 Methods

The study sites, sampling and electrophoretic methods and enzyme polymorphisms have all been described above (Chapters 2 and 3). All complete putative families from which blood samples were obtained were included in the analysis (see Table 4.1 for numbers per site and year).

To avoid the possibility of investigator bias when scoring gels all the blood samples were analysed 'blind' with respect to familial relationships. The unique BTO ring number of a bird was used throughout sample handling to reduce the possibility of identification and labelling errors; blood samples were generally processed in BTO number order (BTO rings were used sequentially and parents were only rarely ringed at the same time as their offspring).

The genetic basis of different enzyme phenotypes has been discussed in Chapter 3. Mismatches amongst progeny phenotypes were detected as those isozyme patterns that appeared to be genetically incompatible with the putative parents, using either of the criteria for the detection of an exclusion as described above. Where mismatches were detected the sample was retested from the original sample tube.

Of the seven loci for which all samples were screened during

Table 4.1

Totals for offspring where both putative parents were sampled, with numbers of clutches in parentheses.

Site	Year			Total
	1980	1981	1982	
Brackenhurst	4	49	62	115
	(1)	(19)	(22)	(42)
Sutton Bonington	24	106	112	242
	(7)	(36)	(39)	(82)
Combined	28	155	174	357
	(8)	(56)	(63)	(124)

this study, only one, EST2, provided good evidence for the presence of null alleles (see Chapter 3). Although exclusion probabilities can be calculated for loci with nulls, as relatively little extra data would be obtained EST2 was omitted from most of the analysis. The remaining protein loci used together with adult gene frequencies are summarised in Table 4.2. Allele frequencies at none of the loci used in the analysis showed significant departure from Hardy-Weinberg expectations (See Chapter 5).

To test the assumptions and validity of calculations concerning probabilities of exclusion and expectations derived from them, a simulation of the effects of non-paternity and non-parentage was conducted by the random reallocation of firstly male parents, then also female parents, to families, and counting mismatches which would then result. This analysis was restricted to the 1981 data for Sutton Bonington (37 families) but all families were included for Brackenhurst (43 families). For the Brackenhurst data, where separate years were included, parents were reallocated within year classes only.

Any copulatory behaviour observed whilst watching nestboxes was recorded in those instances where birds were either known to be associated with a particular nestbox or else were identifiable from colour rings.

4.3 Results

4.3.1 Observed Mismatches

Mismatches were observed to occur in 29 clutches. Type (i) exclusions occurred at all loci; the extra reliability of type (i) exclusions has been referred to above. This category excludes the

Table 4.2

Adult allele frequency estimates for protein loci used in parentage analyses.

Locus	Site	Allele Frequencies						2n
		A	B	C	D	E	O	
6PGD	SB	0.025	0.975					714
	BR	0.033	0.967					486
PEPD3	SB	0.006	0.931		0.063			638
	BR	0.016	0.930	0.002	0.052			442
PEPD2	SB	0.055	0.940		0.006			714
	BR	0.045	0.955					486
IDHC	SB	0.731	0.269					714
	BR	0.691	0.309					486
PEPT	SB	0.003	0.029	0.027	0.937	0.004		714
	BR		0.043	0.031	0.905	0.021		486
GP1	SB	0.161	0.807	0.032				716
	BR	0.183	0.785	0.031				480
EST2*	SB	0.075	0.790	0.008			0.127	716
	BR	0.076	0.837	0.008			0.079	488

* Not used in most parentage analyses; frequencies calculated using maximum - likelihood procedure (see Section 5.3.2).

possibility of being misled by the presence of null alleles, and also guards against the results of other events such as non-disjunction. In some mating combinations, e.g. where one or both parents are heterozygous, an apparent type (ii) mismatch could not result from a hidden parental null allele. Thus in population studies (as opposed to the investigation of individual cases) where the possibility of undetected null alleles is of greatest concern, some exclusions made under criterion (ii) will be as reliable as those made under criterion (i). This equally reliable kind of exclusion, previously regarded as type (ii), is redefined here as type (iii) and described in Table 4.3. The mismatch data obtained here were checked for type (iii) exclusions, but none were present.

The mismatch data are presented in Table 4.4. These include details of all clutches containing affected pulli, the loci at which mismatches occurred, and the exclusion criteria. The frequencies with which the different criteria were used to determine mismatches are summarised in Table 4.5. It should be noted that as null alleles are known to be frequent at the EST2 locus (see Chapter3), mismatches were determined at that locus using criterion (i) only.

A comparison within each site of both the number of nestlings mismatching and the number of clutches containing them (Table 4.6) did not show any significant heterogeneity at any locus among years, or else between sites for all years combined. The number of mismatches at each locus is, however, small, but there are similarly no significant differences in the total proportions of nestlings or clutches mismatching when all loci are considered together.

In most clutches where mismatches were observed only the combination of parental genotypes could be classified as incompatible, but in a minority of cases a specific putative

Table 4.3

Illustration of different mismatch categories for genotypes inferred from phenotypes by making the assumption of codominance. The definitions assume the involvement of a single gene and no polymorphism for gene number.

Minimum number of alleles in population	Mating type	Example	Examples of mismatch type*			Minimum no. alleles to allow type (i) mismatches
			(i)	(ii)	(iii)	
2	incross	AAxAA	AB			2
2	2-way outcross	AAxBB	AC	AA		3
2	backcross	AAxAB	AC	BB		3
2	2-way intercross	ABxAB	AC			3
3	3-way outcross	ABxCC	AD	AA	AB,CC	4
3	3-way intercross	ABxBC	AD		AA,AB, BC,CC	4
4	4-way intercross	ABxCD	AE		AA,AB,BB CC,CD,DD	5

* Definitions:

Type (i) - Neither parent has allele.
(ii) - Expected parental allele lacking (null allele possible).
(iii) - Special case of (ii): not explicable by unseen null.

Table 4.4

Data summary for clutches containing mismatches.

Site	Nest Code	Fledging Date *	Clutch Size	Fledging No.	Number of mismatches per locus							Total† Exclusions	Exclusion‡ Types	Excluded‡ Parent(s)
					6PGD	PEPD3	PEPD2	IDHC	PEPT	GP1	EST2			
BR	G07	0681	3	2						1		1	2	M
	G09	0681	4	2		1					2	2	2,1	M
	G07	0781	4	2		1				1		2	1	-
	G08	0781	4	4						2		2	1	F
	G29	0781	4	3				1		1		1	1,1	-
	G03	0582	3	1					1			1	1	-
	G09	0682	4	4						1		1	2	F
	G12	0682	4	2					1			1	2	M
	G34	0682	4	3	1							1	1	-
	H18	0682	4	4		1						1	1	-
	G33	0882	4	3				2				2	2	F
	Total nestlings			30	1	3	0	3	2	6	2	15		
	Total clutches			11	1	3	0	2	2	5	1	11		

/contd.

Table 4.4 (cont'd)

Site	Nest Code	Fledging Date *	Clutch Size	Fledging No.	Number of mismatches per locus							Total † Exclusions	Exclusion‡ Types	Excluded‡ Parent(s)
					6PGD	PEPD3	PEPD2	IDHC	PEPT	GPl	EST2			
SB	C19	0780	3	3				3				3	2	M
	E02	0681	5	1					1			1	1	-
	V06	0681	4	4						1		1	1	-
	C24	0781	4	3				2				2	1	-
	E21	0781	4	4				4				4	2	M
	E23	0781	5	4						2		2	1	-
	V08	0582	4	3		1						1	1	
	E11	0682	4	3			2					2	1	-
	E12	0682	3	3		1						1	1	-
	E37	0682	4	3		2			1	1		2	1,1,1	-
	E58	0682	4	3				2				2	2	M
	E61	0682	4	3				1				1	1	-
	E66	0682	2	2			1		1			2	1,1	-

/contd.

Table 4.4 (cont'd)

Site	Nest Code	Fledging Date *	Clutch Size	Fledging No.	Number of mismatches per locus							Total† Exclusions	Exclusion‡ Types	Excluded‡ Parent(s)
					6PGD	PEPD3	PEPD2	IDHC	PEPT	GP1	EST2			
SB cont'd														
	C06	0782	4	4				1				1	2	M
	C17	0782	4	4							3	3	1	-
	E03	0782	3	3						1		1	2	F
	E07	0782	4	4					1			1	1	-
	E61	0782	5	2				1				1	1	-
	Total nestlings			56	0	4	3	14	4	5	3	31		
	Total clutches			18	0	3	2	7	4	4	1	18		

* month and year. † some nestlings were excluded at more than locus.

‡ shown respectively for each affected locus (refer to Table 4.3)

‡ M = male incompatible, F = female incompatible.

Table 4.5

Summary of observed exclusion types (refer to Table 4.3). Only one type was observed at any locus within a clutch.

(a) Each locus (clutches)

Site	Exclusion type	Locus 6PGD	PEPD3	PEPD2	IDHC	PEPT	GPI	EST2	Total
BR	(i)	1	1	0	0	1	3	2	8
	(ii)	0	2	0	3	1	3	-	9
SB	(i)	0	4	3	4	4	4	3	22
	(ii)	0	0	0	10	0	1	-	9

(b) Totals (clutches)

Site	Exclusion criterion			
	(i) only	(ii) only	(i) & (ii) (same clutch)	total (i)
BR	6	5	1	7
SB	13	5	0	13
Total		10		20

Table 4.6

Numbers of mismatches observed at each locus, for each year and site. The numbers of clutches containing the mismatches are shown in parentheses.

Site	Year	Locus							* Total mismatches	Sample size
		6PGD	PEPD3	PEPD2	IDHC	PEPT	GP1	EST2		
BR	1980	0	0	0	0	0	0	0	0	4 (1)
	1981	0	2 (2)	0	1	0	5 (4)	2 (1)	8 (5)	44 (19)
	1982	1 (1)	1 (1)	0	2 (1)	2 (2)	1 (1)	0	7 (6)	67 (22)
	Total	1 (1)	3 (3)	0 (0)	3 (2)	2 (2)	6 (5)	2 (1)	15 (11)	115 (42)
SB	1980	0	0	0	3 (1)	0	0	0	3 (1)	24 (7)
	1981	0	0	0	6 (2)	1 (1)	3 (2)	0	10 (5)	108 (36)
	1982	0	4 (3)	3 (2)	5 (4)	3 (3)	2 (2)	3 (1)	18 (12)	110 (39)
	Total	0 (0)	4 (3)	3 (2)	14 (7)	4 (4)	5 (4)	3 (1)	31 (18)	242 (82)

* not necessarily a row total as some individuals will mismatch at more than one locus.

parental genotype could be identified as such. Of course, if a mismatch is due to non-parentage, the assigning of one parent as incompatible does not confirm or rebut parentage by the other parent. This information is also provided in Table 4.4, and summarised below.

Table 4.7

Numbers of clutches and nestlings observed to mismatch with parents of each sex.

	Incompatible parent	
	Male	Female
Clutches containing mismatches	7	4
Nestlings mismatching	14	6

Whilst the number of clutches with incompatible male parents is only marginally greater than the number for incompatible females, the difference is more marked when the total number of mismatching nestlings are considered, but still non-significant ($0.10 > \underline{p} > 0.05$).

4.3.2 Exclusion probabilities

The gene frequencies in the two study populations differed significantly (see Chapter 5; adult gene frequencies are shown in

Table 4.2). Consequently, calculations had to be conducted separately for data concerning each study site. Exclusion probabilities for parentage and for paternity in the single offspring case were calculated using the formulae of Gundel and Reetz (1981), which assume that the population is randomly mating. Additionally, the probability of exclusion for paternity of clutches was calculated for each of the fledging numbers encountered among the families sampled here; these probabilities assume that all fledglings were sired by the same father. The exclusion probabilities for each locus, together with cumulative probabilities, are shown in Table 4.8.

Although there is no strong evidence for deviation from the Hardy-Weinberg equilibrium, deviations of true probabilities of exclusion from those calculated may occur if there are any instances of mating between relatives (Salmon and Brocteur 1975).

4.3.3 Simulation Results

The simulation of non-parentage by the random allocation of parents allows the testing of calculated exclusion probabilities. Indeed, such an approach might be used to generate the probabilities. This has the advantage of basing the probabilities on the actual distribution of multiple locus genotypes of observed putative parents and sibships; the assumptions concerning random mating and the equality of gene frequencies between the sexes are not then necessary. The disadvantages are primarily those of sampling effects, the increased computation required and, with respect to calculations concerning entire clutches, the possibility of inaccuracies due to putative sibships including non-sibs. Putative sibships were kept together here. Any multiple

Table 4.8

Calculated exclusion probabilities based on allele frequencies for entire adult samples.

(a) Probability of excluding incorrectly ascribed pair.

(b) Probability of excluding an incorrect parent assuming one parent is a true parent, for different numbers of offspring (s).

Probability for:	Site	s	LOCUS PGD	PEPD3	PEPD2	IDHC	PEPT	GPl	Combined
(a)	BR	1	0.0562	0.1213	0.0801	0.2580	0.1610	0.3038	0.6694
	SB	1	0.0425	0.0966	0.1023	0.2432	0.1024	0.2302	0.5939
(b)	BR	1	0.0304	0.0666	0.0453	0.2051	0.0888	0.2161	0.5094
		2	0.0438	0.0937	0.0640	0.2494	0.1248	0.2751	0.6137
		3	0.0506	0.1076	0.0736	0.2756	0.1430	0.3077	0.6627
		4	0.0542	0.1149	0.0787	0.2940	0.1526	0.3277	0.6898
		5	0.0561	0.1188	0.0815	0.3067	0.1576	0.3404	0.7057
	SB	1	0.0266	0.0545	0.0588	0.1881	0.0538	0.1492	0.4315
		2	0.0329	0.0768	0.0821	0.2312	0.0770	0.1963	0.5326
		3	0.0381	0.0882	0.0942	0.2563	0.0888	0.2215	0.5809
		4	0.0408	0.0943	0.1008	0.2735	0.0948	0.2362	0.6076
		5	0.0422	0.0977	0.1044	0.2851	0.0980	0.2450	0.6232

paternity/maternity among sibs might consequently bias estimates of non-parentage per clutch. No evidence of this was obtained from the genotypic combinations in sibships, and results concerning parent-offspring trio comparisons should not be affected.

The genotypic frequencies of the subsets of adults being considered were checked for agreement between the sexes and for agreement with the remainder of the adults. There were no significant differences between the sexes within a site and a general agreement with the other adult data in each case. There was no significant difference in heterozygosity between the groups. Exclusion probabilities were recalculated for the subsets (Table 4.9) and found to be close to those previously estimated on the basis of all sampled adults (Table 4.7).

The mismatches resulting from the simulation of non-paternity and non-parentage are shown in Tables 4.10 and 4.11 together with the expectations based upon exclusion probabilities calculated from the entire adult samples. Expectations are not indicated for the different clutch sizes in the non-parentage case as the appropriate formulae have not been devised. In the case of non-paternity it can be seen that the number of simulated clutches containing mismatches was just below the expectation for each data set. The number of trio mismatches was much lower, and appeared to be significantly different from expectation in each case. Trios are inevitably grouped within sibships and it is not known what effect this lack of independence will have upon the probabilities calculated when the data are treated as independent. The combined probability, however, would appear to be very small ($\underline{G} = 11.48$, $\underline{p} < 0.005$), and when the maximum-likelihood distribution resulting when only one offspring per clutch is selected was compared with the expectations

Table 4.9

Calculated exclusion probabilities for parent-offspring trios based on allele frequences in the subsets used for nonparentage simulations.

(a) Probability where both parents are incorrect.
(b) Probability where one parent only is incorrect.

Probability	Subset	Locus						
		6PGD	PEPD3	PEPD2	IDHC	PEPT	GP1	Combined
(a)	SB 1981	0.0133	0.1130	0.0805	0.2070	0.1098	0.2738	0.5874
	BR 1980-82	0.0527	0.1614	0.0794	0.2507	0.1076	0.2841	0.6499
(b)	SB 1981	0.0068	0.0652	0.0456	0.1487	0.0605	0.1776	0.4172
	BR 1980-82	0.0285	0.0922	0.0449	0.1967	0.0584	0.1954	0.4874

Table 4.10. Simulation of nonpaternity by the random reallocation of fathers among families: comparison of results with expecteds based upon calculated nonpaternity exclusion probabilities (P'_s) for each brood size (s). n_s = number of broods of size s.

Sample	Brood size	Clutches	Total Fledgings	Expected * excluded proportion of clutches	Excluded clutches *		Mismatch trios		Mean number mismatches per excluded clutch	
					observed	expected	observed	expected	observed	expected
	s	n_s	$n_s s$	P'_s	-	$n_s P'_s$	-	$n_s s P'_1$	-	sP'_1/P'_s
SB 1981	1	2	2	0.4315	1	0.86	1	0.86	(1)	(1)
	2	6	12	0.5326	2	3.20	2	5.18	1.00	1.62
	3	17	51	0.5809	9	9.88	20	22.00	2.22	2.23
	4	12	48	0.6076	6	7.29	14	20.71	2.33	2.84
Total		37	113		18	21.22	37	48.75		
%					(48.6%)	(57.4%)	(32.7%)	(43.2%)		
Goodness of fit test					G = 1.13 n.s.		G = 5.12 $p < 0.05$			

BR
1980 cont'd on next page.
-82

Table 4.10 cont'd

Sample	Brood size	Clutches	Total Fledglings	Expected * excluded proportion of clutches	Excluded clutches *		Mismatch trios		Mean number mismatches per excluded clutch	
					observed	expected	observed	expected	observed	expected
	s	n_s	$n_s s$	P_s	-	$n_s P_s$	-	$n_s s P_1$	-	sP_1/P_s
BR 1980 -82	1	3	3	0.5094	2	1.53	2	1.53	(1)	(1)
	2	15	30	0.6137	11	9.21	16	15.28	1.45	1.66
	3	12	36	0.6627	6	7.95	12	18.34	2.00	2.31
	4	12	48	0.6898	6	8.28	16	24.45	2.67	2.95
Total		42	117		25	26.96	46	59.60		
%					(59.5%)	(64.2%)	(39.3%)	(50.9%)		
Goodness of fit test					G = 0.39 n.s.		G = 6.36 $p < 0.05$			

* an excluded clutch is one which contains a mismatch.

Table 4.11 Simulation of nonparentage by the random reallocation of fathers among families: comparison of results with expecteds based upon the calculated nonparentage exclusion probability (P''_1) for a single offspring.

Sample	Brood size	Clutches	Total fledglings	Expected * excluded proportion	Excluded clutches *		Mismatch trios		Mean number mismatches per excluded clutch
					observed	expected	observed	expected	observed
	s	n_s	$n_s s$	P''_1	-	$n_s P''_1$	-	$n_s s P''_1$	-
SB 1981	1	2	2	0.5939	1	1.19	1	1.19	(1)
	2	6	12		3		6	7.13	1.17
	3	17	51		12		31	30.29	2.58
	4	12	48		7		16	28.51	2.29
	Total	37	113		23		54	67.11	
	%				(62.2%)		(47.8%)	(59.4%)	
	Goodness of fit test						G = 6.18	$p < 0.05$	

Table 4.11 cont'd

Sample	Brood size	Clutches	Total fledglings	Expected * excluded proportion	Excluded clutches *		Mismatch trios		Mean number mismatches per excluded clutch
					observed	expected	observed	expected	observed
	s	n_s	$n_s s$	P_1	-	$n_s P_1$	-	$n_s s P_1$	-
BR 1980 -82	1	3	3	0.6694	2	1.34	2	1.34	(1)
	2	15	30		12		18	20.08	1.5
	3	12	36		10		25	24.10	2.5
	4	12	48		9		22	32.13	2.44
	Total	42	117		33		67	78.32	
	%				(78.6%)		(57.3%)	(66.9%)	
	Goodness of fit test						G = 4.75	$p < 0.05$	

* an excluded clutch is one which contains a mismatch.

(a very inefficient test) the result for the combined data remained significant ($\underline{G} = 4.57$, $\underline{p} < 0.05$). (The maximum likelihood values for single independent offspring were found by taking the mean number of mismatches per individual within each clutch.) A similar deficiency of mismatching trios occurred in the simulation of non-parentage (Table 4.11). Consequently, the assumptions concerning exclusion probabilities and their application here would be seen to be in some respect inadequate.

The results obtained could be explained if the randomly reallocated parental genotypes (considering all loci together) bore a closer relationship to the actual parent than random expectation. There is only slight evidence for any gametic phase disequilibria , however, and no analyses of higher order disequilibria (Weir 1979) have been carried out. Whilst the reallocation of parents to families might occasionally result in an incorrect putative parent being replaced by the true parent, this will on average be very improbable. The observed deficiency of mismatches might, however, result if the samples contain groups of closely related individuals. If this were the source of discrepancy from expectation then the best estimate of an exclusion probability would depend very much on the spatial distribution and behaviour of related birds.

4.3.4 Multiple-locus Mismatching

In an attempt to gain a further insight into the determination of exclusions the distribution of mismatches among the 6 loci used was investigated in both the actual and simulation data.

The probability of detecting an incorrect offspring that exhibits a total of $\underline{M}$ mismatches out of $\underline{N}$ loci tested may be determined as the sum for all combinations of loci of:

{probability of observing no mismatches at ($\underline{N}$ - $\underline{M}$) loci combined} x {probability of observing mismatches at all remaining $\underline{M}$ loci combined}.

Hence, the discrete probability distribution for the number of loci at which an individual incorrect offspring is expected to mismatch is given by the terms of the series:

$$\prod_{i=1}^{i=N}(1-P_i) \;,\quad \sum_{j=1}^{j=N}\Big(\Big(\prod_{\substack{i=1\\ i\neq j}}^{i=N}(1-P_i)\Big)P_j\Big) \;,\quad \sum_{k=1}^{k=N}\sum_{\substack{j=1\\ j<k}}^{j=N}\Big(\Big(\prod_{\substack{i=1\\ i\neq j\\ i\neq k}}^{i=N}(1-P_i)\Big)P_jP_k\Big) \;,$$

$$\sum_{l=1}^{l=N}\sum_{\substack{k=1\\ k<l}}^{k=N}\sum_{\substack{j=1\\ j<k}}^{j=N}\Big(\Big(\prod_{\substack{i=1\\ i\neq j\\ i\neq k\\ i\neq l}}^{i=N}(1-P_i)\Big)P_jP_kP_l\Big) \;,$$

$$\sum_{m=1}^{m=N}\sum_{\substack{l=1\\ l<m}}^{l=N}\sum_{\substack{k=1\\ k<l}}^{k=N}\sum_{\substack{j=1\\ j<k}}^{j=N}\Big(\Big(\prod_{\substack{i=1\\ i\neq j\\ i\neq k\\ i\neq l\\ i\neq m}}^{i=N}(1-P_i)\Big)P_jP_kP_lP_m\Big) \;,$$

$$\sum_{n=1}^{n=N}\sum_{\substack{m=1\\ m<n}}^{m=N}\sum_{\substack{l=1\\ l<m}}^{l=N}\sum_{\substack{k=1\\ k<l}}^{k=N}\sum_{\substack{j=1\\ j<k}}^{j=N}\Big(\Big(\prod_{\substack{i=1\\ i\neq j\\ i\neq k\\ i\neq l\\ i\neq m\\ i\neq n}}^{i=N}(1-P_i)\Big)P_jP_kP_lP_mP_n\Big) \;,\;\ldots\ldots\;,\prod_{i=1}^{i=N}P_i$$

where $\underline{N}$ is the total number of loci and $\underline{P_i}$ is the probability of exclusion at the $\underline{i}$th locus. Appropriate terms for the case where

$\underline{N}$ = 6, as in this analysis, are shown in full.

The expected proportions of each mismatch class were calculated for exclusion probabilities based on the simulation subset. From Figure 4.1, it can be seen that the observed and expected distributions are rather different, with a large excess of non-excluded offspring in 3 of 4 cases. (Note that the comparisons involving low expecteds were excluded from the goodness of fit $\underline{G}$ tests - refer to Section 5.2.) The allowance which should be made for non-independence within clutches, as referred to above, when interpreting goodness of fit $\underline{G}$-statistics is again unknown. The raw values for $\underline{G}$ are, however, very large and are in each case accounted for principally by the excess of non-exclusions. If, as previously, allowance for non-independence is made by taking maximum-likelihood values for the distribution that would result by randomly selecting one offspring per clutch, then the non-exclusions are found to persist in every case, but only significantly in two of four (Table 4.12). However, when only those nestlings which contain mismatches are considered and expected distributions calculated (Figure 4.2), it can be seen that there is no consistent deviation from expectation. Without allowance for non-independence, 2 of the 4 $\underline{G}$ values are large, but each is due to opposite deviations from expectation. When, as before, the rather inefficient maximum-likelihood estimates for independent cases are considered (Table 4.13), none of the observed distributions deviates significantly from expectation.

Hence the number of mismatches occurring in an individual excluded case is, on average, close to expectation. Thus it appears that the deficiency of exclusions in the simulation results is due to an excess of totally non-excluded cases, rather than to a general

Figure 4.1

Comparison of observed and expected mismatches per individual (for 6 loci) in simulations of (a) nonpaternity and (b) nonparentage, for all nestlings in subsamples: (i) SB 1981, (ii) BR 1980-82. ▨ = observed, □ = expected.

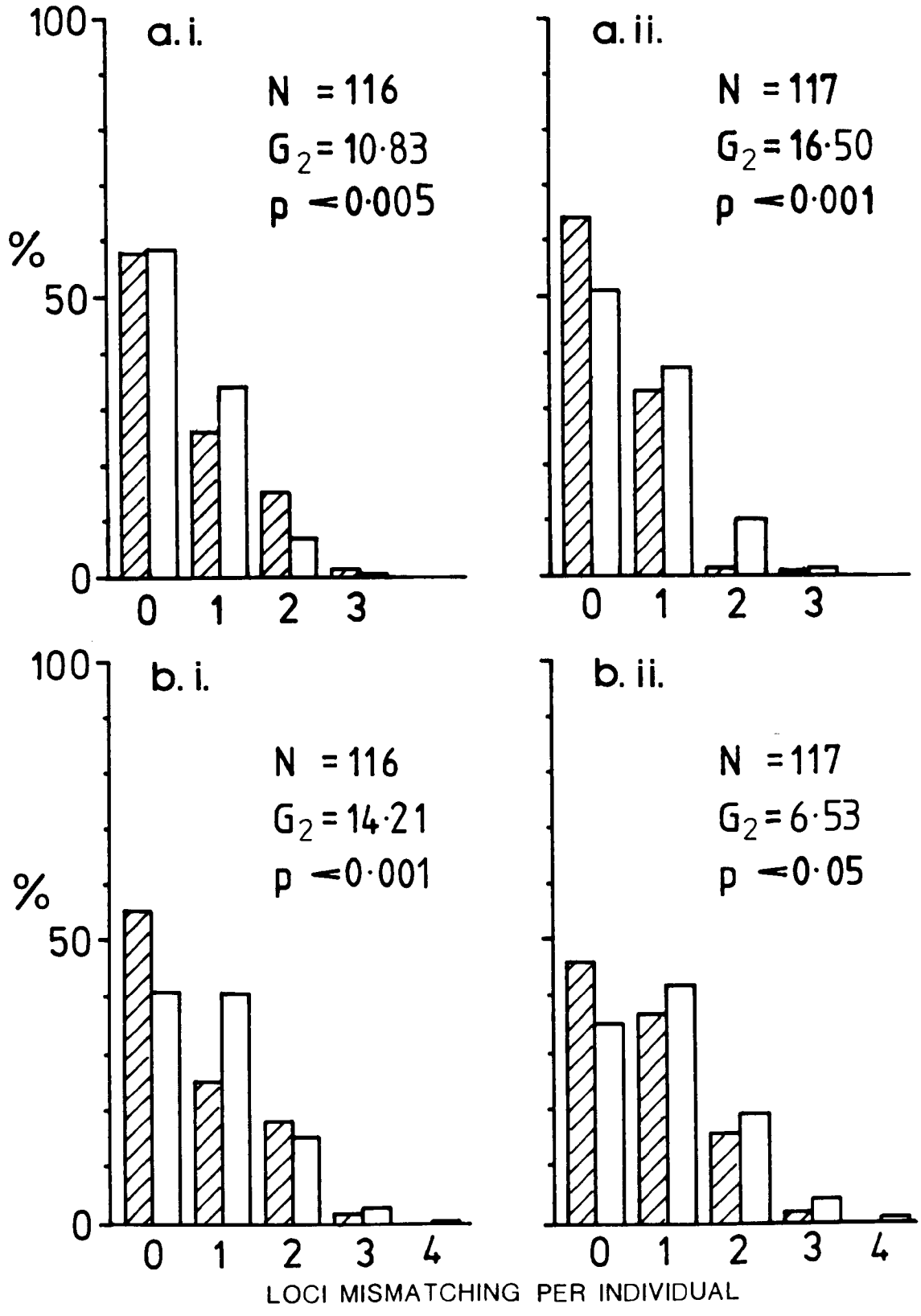


Figure 4.2

Comparison of observed and expected mismatches per individual (for 6 loci) in simulations of (a) nonpaternity and (b) nonparentage, for nestlings with at least one mismatch: (i) SB 1981 subsample, (ii) BR 1980-82 subsample. ▨ = observed, □ = expected.

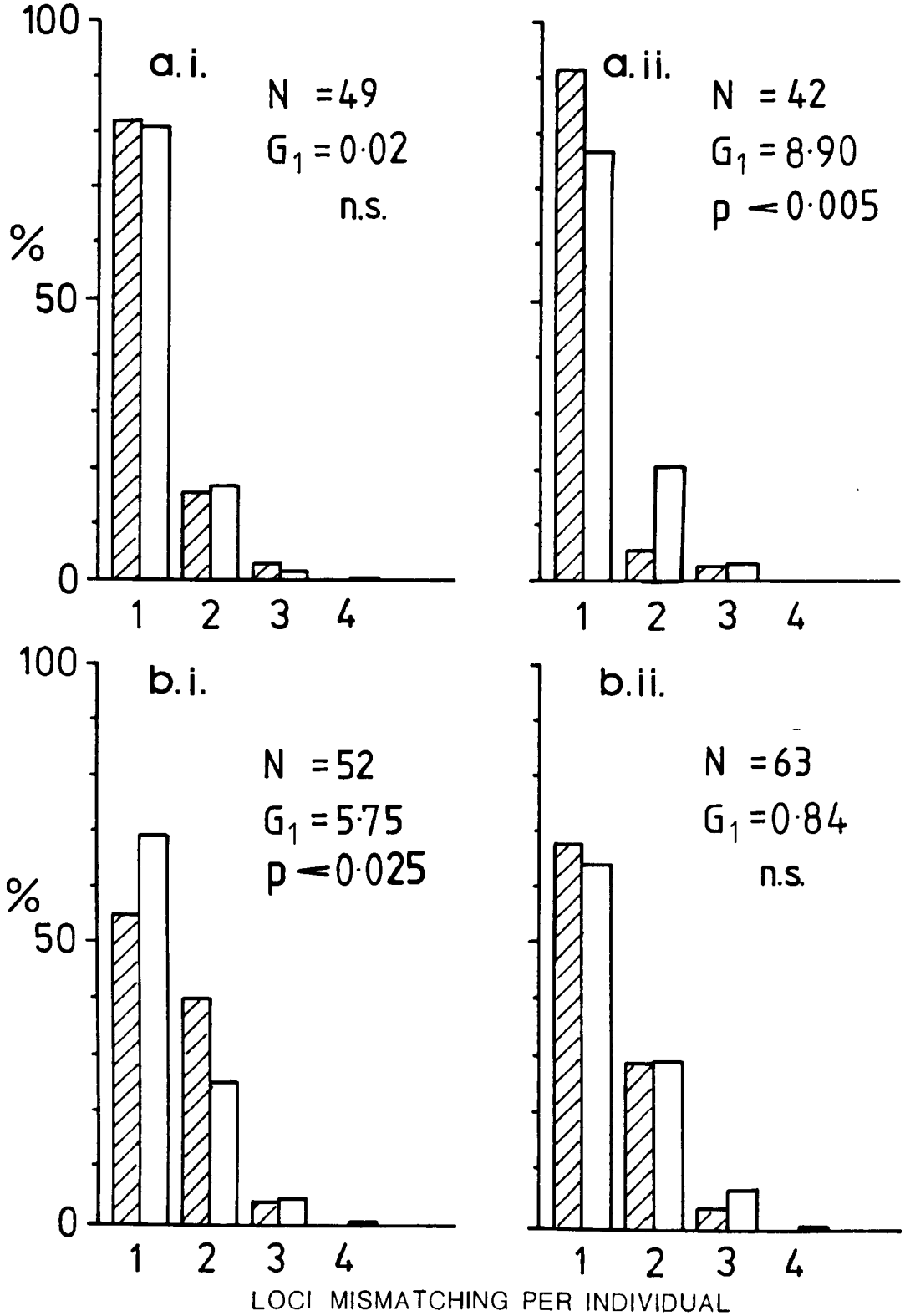


Table 4.12

Simulations of (a) nonpaternity, and (b) nonparentage: estimated frequencies of the number of loci mismatching in independently selected offspring, compared with expected frequencies. Refer to text for details of selection.

Simulation	Subset	Number of selected offspring	Frequency	Mismatches per offspring				Goodness of fit
				0	1	2	3	G
(a)	SB 1981	37	obs	24.3	10.3	2.0	0.3	0.95
			exp	21.6	12.5	2.7	0.3	
	BR 1980 -82	42	obs	25.3	15.5	0.8	0.5	5.00*
			exp	21.5	15.7	4.2	0.5	
(b)	SB 1981	37	obs	19.1	9.1	7.9	0.8	4.49*
			exp	15.3	15.1	5.6	1.0	
	BR 1980 -82	42	obs	17.8	15.9	7.5	0.8	3.03
			exp	14.7	17.4	7.9	1.7	

* $p < 0.05$

Table 4.13

Simulations of (a) nonpaternity, and (b) nonparentage: estimated frequencies of the number of loci mismatching in those independently selected offspring which contain mismatches, compared with expected frequencies. Refer to text for details of selection.

Simulation	Subset	Number of selected offspring with mismatches	Frequency	Mismatches per offspring			Goodness of fit
				1	2	3	
(a)	SB 1981	12.7	obs	10.3	2.0	0.3	0.09
			exp	10.2	2.2	0.2	
(b)	BR 1980 -82	16.8	obs	15.5	0.8	0.5	3.71
			exp	12.8	3.5	0.4	
(a)	SB 1981	17.8	obs	9.1	7.9	0.8	3.16
			exp	12.4	4.6	0.8	
(b)	BR 1980 -82	24.3	obs	15.9	7.5	0.8	0.87
			exp	15.5	7.0	1.6	

deficiency of mismatches spread uniformly amongst cases. This result could again be explained by the occurrence of related birds within the randomly reallocated families.

4.3.5 Distributions of Actual Mismatches.

In the simulation described above the probability of reallocated parents being true parents was extremely low. Hence the simulated putative families were known to be incorrect and the numbers of the different categories of non-exclusion and exclusion could all be compared with expectation. With regard to the data collected for actual putative families the absolute number of incorrect families is, of course, not known directly and the various expected distributions can be estimated only from the observed mismatches.

(a) Distribution Amongst Loci.

The frequency of mismatches observed using a particular marker locus is expected to equal the exclusion probability for that locus. Which of the calculated probabilities will be more appropriate depends upon the nature of the exclusion events (see above). The possible inadequacies of calculated exclusion probabilities and the assumptions underlying them have been referred to already. Whatever might be the exact values for the probabilities, their relative values should reflect those calculated. If $\underline{I}$ is the proportion of actually incorrect putative offspring in a tested sample then the proportion of mismatches expected to be observed at locus $\underline{i}$ is given by $\underline{M}_i = \underline{I}\underline{P}_i$, where $\underline{P}_i$ is the probability of exclusion at locus $\underline{i}$. Thus a regression of $\underline{M}_i$ against $\underline{P}_i$ should provide a measure of the agreement among loci of estimates of $\underline{I}$ ($\hat{\underline{I}}$)

and, from the gradient, an overall value for $\hat{I}$ according to the assumptions on which the particular exclusion probabilities used were based.

Regressions were therefore performed for the data from each site using exclusion probabilities based separately on hypotheses of non-parentage and non-paternity. Regressions were also performed for data from both sites combined, an analysis relying on the assumption that I will have a similar value at both sites. The results are given in Table 4.14, where it can be seen that the gradients are significant in every case and especially so ($p = 0.001$) for the combined data (Figures 4.3 and 4.4). The intercepts are always extremely close to zero and non-significant, agreeing with the expectation of a regression of the form $Y = bX$ (or $M_i = IP_i$). Thus there is good agreement among the marker loci concerning the value of I with the estimate $\hat{I}$ depending on assumptions concerning the origin of exclusions.

Without correction for any errors (see Section 4.4) the 95% confidence estimates for $\hat{I}$, based on the combined samples, would appear to lie between 9.9-24.1% if incorrect attribution of both parents is the source of exclusions and between 12.1-29.7% if it is the incorrect attribution of a single parent. Again, there is some lack of independence amongst the data (on average for each locus considered separately, the mean number of mismatches contributed to the analysis by a clutch containing mismatches was 1.36). Allowance for this would extend the confidence limits for $\hat{I}$.

Table 4.14

Summary of regressions on exclusion probability for trios assuming (a) both parents are incorrect or (b) only one parent is incorrect. Refer to Figures 4.3 and 4.4.

Assumption	Sample	N	intercept	slope	±	S.E.	probability for slope
(a)	SB	30	-0.006	0.197	±	0.070	0.047
	BR	15	-0.004	0.158	±	0.046	0.026
	Combined	45	0.004	0.170	±	0.036	0.001
(b)	SB	30	-0.002	0.264	±	0.072	0.022
	BR	15	0.002	0.181	±	0.064	0.048
	Combined	45	0.001	0.209	±	0.045	0.001

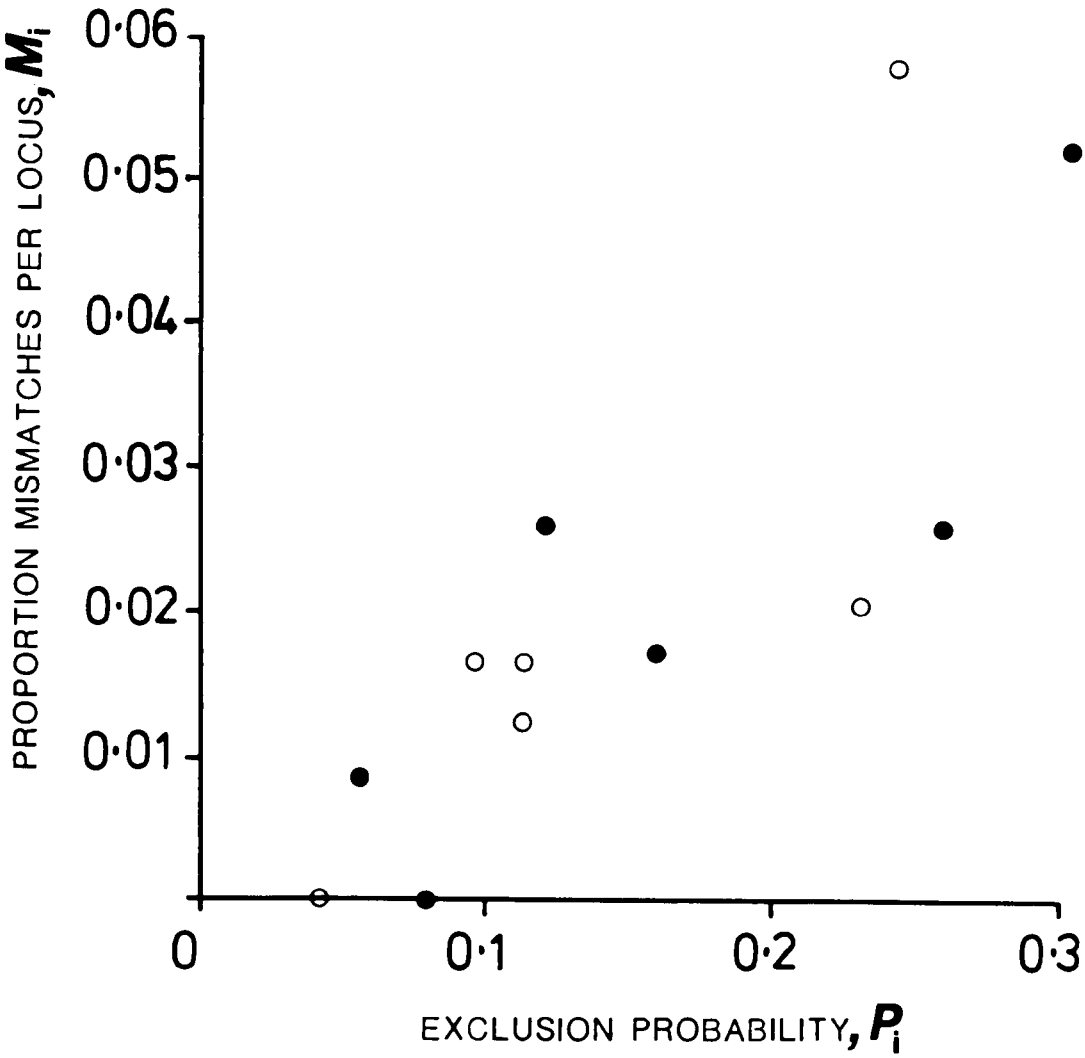


Figure 4.3

Plot of the proportion (M_i) of nestlings mismatching for locus i against the probability (P_i) of excluding incorrect parentage. Refer to Table 4.14(a) and Section 4.3.5(a).

O = SB81 subsample, ● = BR80-82 subsample.

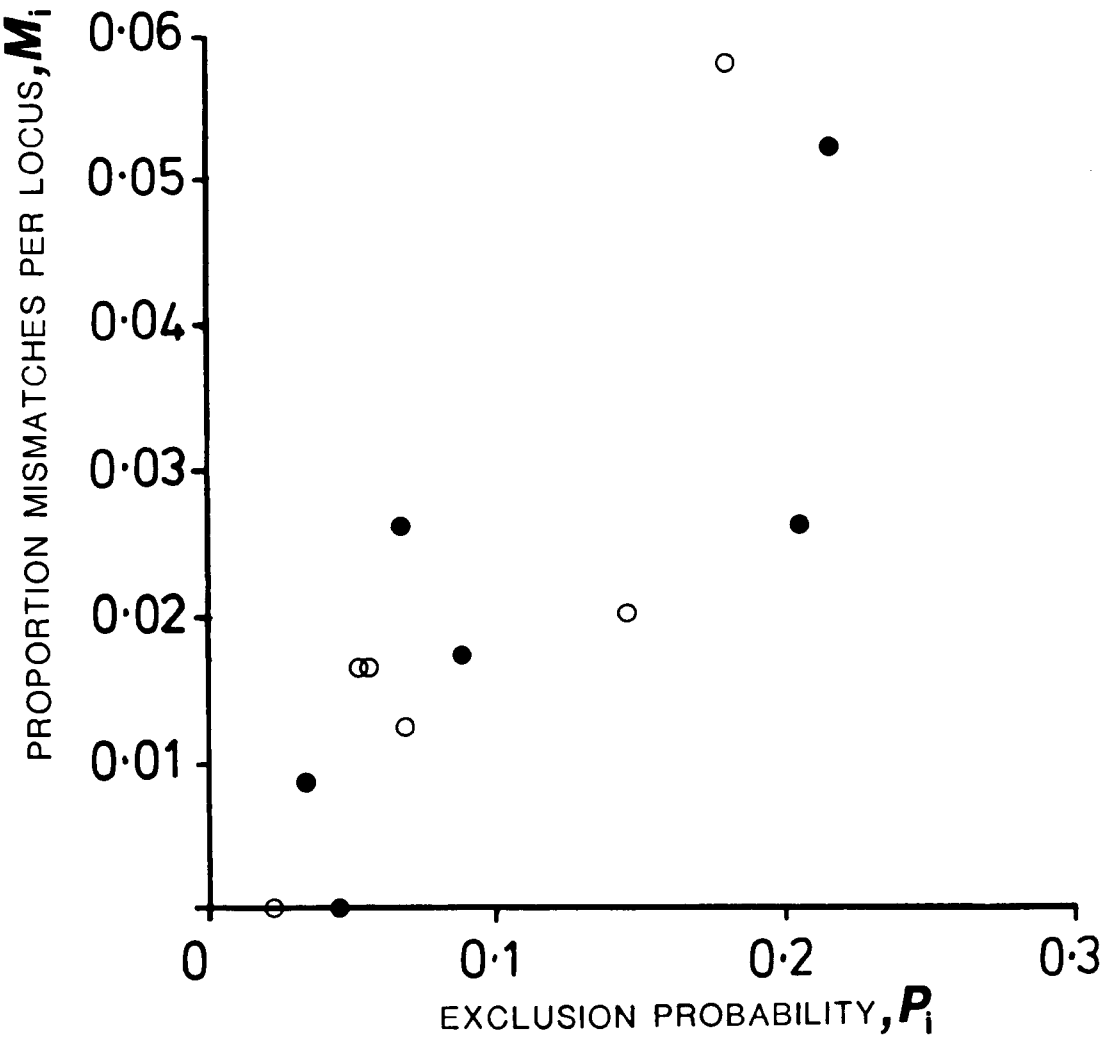


Figure 4.4

Plot of the proportion (M_i) of nestlings mismatching for locus i against the probability (P_i) of excluding incorrect paternity. Refer to Table 4.14(b) and Section 4.3.5(a).

O = SB81 subsample, ● = BR80-82 subsample.

(b) Distribution of Exclusions Among Clutch Sizes.

It is desirable to know whether only a proportion of a clutch is incorrect or whether the entire clutch is affected. For example, if mismatches are a consequence of egg-dumping then only a proportion of the clutch is expected to be affected, whilst if mismatches result from extrapair copulations then an entire clutch might be affected.

The proportion of a clutch that is expected to be excluded in clutches containing at least one exclusion depends on the clutch size; these expected values are known for the case of exclusions resulting from non-paternity, providing that it is assumed that all offspring are sired by the same male (Table 4.10). It was found that the number of exclusions within clutches was in all comparisons lower than expected, and significantly so ($p < 0.05$) for clutch sizes of 3 or more at Sutton Bonington and 4 at Brackenhurst (Table 4.15).

The mean values for different clutch sizes at fledging indicate that the proportion of incorrect offspring types decreases with clutch size and the ratio of observed to expected also decreases (Table 4.16). In the simulations described above, where all offspring per clutch would be incorrect with respect to paternity, it was found that a lower than expected overall detection rate did not result in a decreased number of exclusions within those clutches actually excluded. Thus the results here suggest that for clutches actually observed to contain mismatches only a proportion of the offspring within a putative sibship will, on average, be incorrect, and that this proportion is usually smaller in larger clutches. Indeed, there is only a slight increase in the mean number of observed mismatches with clutch size (actually none at

Table 4.15

The number of mismatching offspring per clutch containing mismatches, compared with the number expected if one putative parent were incorrect for the entire clutch.

Site	Clutch size †	No of clutches	Number of nestlings ‡	M	C	Goodness of fit G	Binomial probability (1-tailed)
SB	2	2	obs.	3	1	-	
			exp.	3.2	0.8		
	3	9	obs.	15	12		
			exp.	20.1	6.9	4.44*	
	4	5	obs.	9	11		
			exp.	14.2	5.8	5.87*	
BR	2	3	obs.	4	2		
			exp.	4.98	1.12	0.57	
	3	3	obs.	4	5		0.033
			exp.	6.93	2.07		
	4	3	obs.	4	8		
			exp.	8.85	3.15		0.004

† at fledging
‡ M - mismatching C - compatible
* $p < 0.05$

Table 4.16

The proportion of mismatching offspring per clutch containing mismatches, relative to the number expected if one putative parent were incorrect for the entire clutch, for clutches of different size.

Site	Clutch size *		
	2	3	4
SB	0.93	0.75	0.63
BR	0.80	0.58	0.45

* at fledging

Brackenhurst) and for all clutches of 3 or 4 (most of those affected) there is a mode of 1 excluded offspring.

The expected distribution for the number of exclusions within excluded clutches of a particular size is unknown - it is certainly not a simple binomial - but the expected means calculated from exclusion probabilities for paternity on the assumption that the entire clutch is affected would require a mode of at least 2 exclusions in all cases, even where only two offspring are incorrect. (If both parents were incorrect then an even higher mode would be expected.) Reference to Table 4.4, however, reveals that the number of exclusions per excluded clutch varies widely, with several instances where the entire clutch is affected. Therefore, there would appear to be many instances where exclusions are the result of single incorrect offspring within clutches (as indicated by a mode of one mismatch), but others where several or all of the clutch are affected (as evidenced by individual cases and means of value exceeding 1).

4.3.6 Field Observations

Any data concerning the direct observation of copulations and/or egg laying are obviously of relevance to this analysis. Few copulations were actually observed where both individuals could be positively identified. As clutches were usually examined at 2 or 4 day intervals, only a proportion of any eggs that were 'dumped' would be detectable. Fluctuations from a presumed laying schedule of one egg per day were only occasionally observed and explicable by dumping. These observations will be discussed further below.

4.4 Discussion

4.4.1 Are Mismatches Error-based or Behaviour-based?

Ashton (1981) distinguished between 'error-based' and 'behaviour-based' mismatches, i.e. those resulting from errors in sample recording, handling and gel scoring, and those resulting directly from the behaviour of the tested individual. Such errors will, of course, increase estimates of non-parentage. That errors will occur in studies of this kind has been widely acknowledged. For example, the requirement of some forensic investigators in human paternity cases of demonstrating particular classes of exclusion and/or at least two exclusions per case has already been referred to. For population studies in humans at least two studies have attempted to estimate the rate of error-based mismatches (Ashton 1980, Lathrop et al. 1983), whilst the others referred to above have not. None of the studies of animal populations using electrophoretic markers for parentage analyses (see above) appear to have allowed for the possibility or effects of errors.

Testing for Error-based Mismatches

Errors fall into two principal categories: field and laboratory. Field errors are those which involve the mis-identification of individuals and mislabelling of tissue samples; they may effect both single and double parent comparisons. Depending on the assumptions made concerning possible parentage the occurrence of field errors may not be quantifiable as any resulting mismatches will mimic possible behavioural events. Laboratory errors, however, are expected to result in mismatches at individual tested loci only, as marker loci are generally scored and recorded

independently. The predicted unique probability distribution for the number of loci expected to mismatch for incorrect parent-offspring comparisons has been elucidated here (see above), and may be used to estimate laboratory errors by comparison with the observed distribution. A similar approach has been used by Ashton (1981), but using a Poisson expected distribution (which fitted reasonably well with the observed distribution for his control sample, but using many more loci than here). Intuitively, if laboratory errors occur then only one locus is likely to be affected for any particular individual. Hence Foltz (1981), in his assessment of extrapair copulations in Peromyscus polionotus was reassured by the occurrence of cases having two loci incompatible, but unfortunately presents insufficient data to allow a check with the expected occurrence of multiple exclusions, and otherwise assumes no laboratory error to have occurred.

In the present study the 46 observed incompatible trios included only 4 instances of double mismatches. Expected distributions (see Fig. 4.2) are shown for comparison in Table 4.17. For expectations based upon single parent incompatibilities, the deficiency of multiple mismatches within sites was not significant, but this probably reflected the small sample sizes as combining the data produced a very significant result ($G = 8.1$, $p < 0.005$). The deficiency with respect to expecteds based upon double parent incompatibilities was inevitably larger, and significant even within sites.

Thus it must be concluded that laboratory errors have occurred here. This would appear to have happened despite stringent handling and checking procedures. As described above (Section 4.2) the genotypes of individuals within trios which mismatched were

Table 4.17

Number of mismatches per excluded trio, m, compared with the distribution expected for (i) non-paternity and (ii) non-parentage. Goodness of fit has been estimated by pooling the classes of two or more mismatches.

Site		m				Goodness of fit
		1	2	3	4	
SB	Observed	26	2	0	0	
	Expected (i)	22.6	4.8	0.5	0.0	n.s.
	Expected (ii)	19.5	7.2	1.3	0.1	$G_1 = 9.13$, $p < 0.05$
BR	Observed	12	1	0	0	
	Expected (i)	9.9	2.7	0.4	0.0	n.s.
	Expected (ii)	8.3	3.8	0.8	0.1	$G_1 = 5.75$, $p < 0.01$

rechecked from the original sample tube in which blood samples were first collected. If these samples were misidentified at the time of collection then no deficiency of multiple exclusions would result as all loci should be affected. Thus the genetic interpretation of actual, repeatable gel patterns for particular loci in some individual samples would appear to be questionable. In addition to the distinctive and frequently reversible kinds of modification already discussed, the possibility of unnoticed modifications resulting from the actions of the gene products of other, possibly variable, loci cannot be ruled out. In conclusion, whilst it must be accepted that laboratory errors have occurred, the source of these errors remains unknown, but post-translational modification is suggested as the least unlikely. Potential sources of laboratory error and the reliability of field techniques used are discussed further below.

4.4.2 Sources of Laboratory Error

A fuller account of the enzyme systems and interpretation of gel patterns is given elsewhere (Chapter 3). Events whose incorrect genetic interpretation could result in mismatches include unidentified null alleles, mutation, unusual recombination events, unidentified gene number polymorphisms and post-translational modifications to the visualised proteins.

Studies of Drosophila (Voelker et al. 1980, Langley et al. 1981) and the conveniently haploid, meiotically derived megametophytes of Pinus spp. (Allendorf, Knudsen and Blake 1982) suggest that many, if not most, electrophoretic loci will have some null alleles (with mean frequency of the order of 0.003). It has been shown that unidentified null alleles could not have been the

cause of most exclusions (Section 4.3.1 and Table 4.5). During this study direct evidence for the presence of null alleles was obtained only at the EST2 locus which was for that reason largely omitted from the mismatch analyses described. No null homozygotes were observed at any of the other loci and no significant deviation from Hardy Weinberg expectations made on the assumption of no null alleles was found (Chapter 5).

The possibility that null alleles may nevertheless account for a significant proportion of mismatches must, however, be considered. The theoretical maximum frequency (with 95% confidence) of a null allele that might remain undetected through the non-appearance of null homozygotes, assuming no selection, can be estimated and in the samples taken here might in an extreme case approach 7-8% for a single locus. Smith (1970) has shown that null alleles will in general cause significant deviation from Hardy-Weinberg expectation only when they are frequent enough for null homozygotes to occur (assuming there is no selection against them). At such a high frequency the resulting excess of incorrectly classified homozygotes would, however, be suspiciously large even if not statistically significant. Null alleles are particularly characteristic of esterase loci. There is no reason to expect the overall frequency of nulls to be substantially different from that observed in other organisms. A generalised formula for the frequency of exclusions (P_{-X}) expected to result from the presence of hidden null alleles has been derived (Table 4.18) as follows:

$$P_X = p_o \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^{n} p_i p_j (3p_i/2 + p_o + \sum_{\substack{k=1 \\ k \neq i \\ k \neq j}}^{n} p_k)$$

where p_o is the frequency of the null allele and p_i is the frequency at the ith of n detectable alleles. For any particular value of N, P_X is maximised when the known alleles are at equal frequency. The value of P_X has therefore been estimated for the most intermediate allele frequencies observed here. Taking p_o = 0.005 as a conservative high estimate for the mean null frequency, the maximum value for P_X obtained was 0.16% (for both IDHC and GP1 frequencies at Brackenhurst). p_o may of course exceed 0.005 for any individual locus, but even if p_o is increased to 0.01 the value of P_X (0.31%) remains small. The rate of exclusions at the same loci was 2-5% of tested trios (Figure 4.3). Therefore it may be concluded that unidentified null alleles were not a significant cause of error-based mismatches in this study.

Of the other possible causes of error-based mismatches, mutation and rare chromosomal rearrangements are even less likely than concealed nulls. Nei (1975) estimates the frequency of mutations that will result in electrophoretically distinguishable gene products to be between 10^{-6} and 10^{-7} per locus per year, while chromosomal events which might do the same are believed to occur at a frequency of about 10^{-5} per locus per year. Gene number polymorphisms have been documented for the haemoglobin alpha chain locus in humans (Rucknagel and Winter 1974) and an amylase locus in the bank vole, Clethrionomys glareola (Nielsen 1977). It would

Table 4.18

Derivation of the probability of the non-detection of a null allele class leading to the wrongful exclusion of a parent-offspring trio.

Mating type	Generalised examples*	Expected frequency*	Accidentally excluded offspring genotypes	Proportion excluded	Resulting probability of exclusion
3-way intercross	A_iA_o x A_iA_j	$2p_o\sum_{\substack{i=1\\j=1}}^{n} p_i^2 p_j$	A_jA_o	1/4	$(p_o\sum_{\substack{i=1\\j=1}}^{n} p_i^2 p_j)/2$
	A_iA_o x A_jA_o	$2p_o^2\sum_{\substack{i=1\\j=1}}^{n} p_i p_j$	A_iA_o, A_jA_o	1/2	$p_o^2\sum_{\substack{i=1\\j=1}}^{n} p_i p_j$
3-way outcross	A_iA_o x A_jA_j	$2p_o\sum_{\substack{i=1\\j=1}}^{n} p_i p_j$	A_jA_o	1/2	$p_o\sum_{\substack{i=1\\j=1}}^{n} p_i p_j^2$
4-way outcross	A_iA_o x A_jA_k	$2p_o^2\sum_{\substack{i=1\\j=1\\k=1}}^{n} p_i p_j$	A_iA_o, A_jA_o	1/2	$p_o\sum_{\substack{i=1\\j=1\\k=1}}^{n} p_i p_j p_k$
TOTAL				$p_o\sum_{\substack{i=1\\j=1\\i\neq j}}^{n} p_i p_j (3p_i/2 + p_o + \sum_{\substack{k=1\\k\neq i\\k\neq j}}^{n} p_k)$	

For the special case of 2 alleles (excluding nulls), this reduces to:

$$\{pqs(3 + s)\}/2$$

where s is the frequency of the null allele class.

* A_o = null allele; p_o = null allele frequency; n = number of alleles (excluding nulls) for the locus; $i \neq j \neq k$.

appear unlikely that such a mechanism could explain mismatches at even one of the loci used, and there was certainly no direct evidence (in the form of individuals with three or more allozymes from a single locus) for this type of polymorphism. The same arguments would apply to the possibility of variable duplicated loci.

There remains the possibility of post-translational modification of isozymes, and this would seem to be the most likely explanation of the mismatches. Observed modifications, and the methods used in some instances to rectify them, have been previously discussed with respect to the loci investigated in this study (Chapter 3).

Some reversible post-translational modification of enzymes was identified at the peptidase and esterase loci and samples were treated appropriately (Chapters 2 and 3). The possibility that some unidentified forms of modification affected individual samples cannot be excluded, however, and would seem the most likely explanation for the excess mismatches. No breeding data could, unfortunately, be obtained from isolated pairs: pairs kept in cages in the laboratory could not be induced to breed. House sparrows will breed in captivity if kept in large aviaries (Washington 1973, Mitchell and Hayes 1973) but this was not possible during this study. The deviations from expected Mendelian inheritance that might result from modification of allozymes or the presence of null alleles that would be required to produce the observed levels of mismatches would in any case be low; very large breeding colonies would therefore be required to ensure their detection.

4.4.3 Reliability of Field Techniques

The effects of errors made in the field will generally, unfortunately, be indistinguishable from those of the sampled individuals' behaviour. Procedures used in the field were designed to minimise the opportunities for mistakes to be made and are described elsewhere (Chapter 2). It is important to check, as far as possible, that no field errors have occurred. The main sources of errors will be the mislabelling of samples and the misidentification of putative parents. Considering the former, rings were used sequentially and tubes labelled with the ring numbers. It is to be expected, therefore, that mismatches occurring due to mislabelling would be clumped within time periods and/or sites, but there was no evidence of this (Section 4.3.1).

The procedures used to identify parents, and if necessary trap them, have been described (Chapter 2). As a check that the potentially less reliable procedures for ascertaining the identity of putative parents will not have been a source of mismatches and hence exclusions, a comparison of the methods has been made between clutches that contained mismatches and those that did not. The numbers of parents identified by the different methods are shown in Table 4.19, where it can be clearly seen that there was no difference between the two groups. As it was not possible to determine specifically which mismatches were not due to laboratory error (with the exception perhaps of the few clutches with mismatches at more than one locus), it was unfortunately not possible to restrict this comparison of the reliability of the method of identification to the group for which the possibility of field errors was of most significance.

Table 4.19

A comparison of the reliability of the identification of the parents of a clutch between clutches with and without mismatches.

Three categories are used for the reliability of parental identification:

(i) High. **Both** parents were positively identified on two or more feeding visits to the clutch or, if only recorded on one visit, were also identified at an earlier or later clutch at the same nest.

(ii) Intermediate. Identification was made either on the basis of:

(a) one feeding visit,

or (b) other activity at the nest as well as at a later or earlier clutch at the same nest.

(iii) Low. Identification was due to activities at the nest other than feeding, or by trapping inside the nestbox.

Reliability category of least certain parent of a clutch	Number of clutches per category[1]		
	Mismatches observed		No mismatches
	m > 1	m = 1	m = 0
High reliability	2	9	33
Intermediate reliability	2	10	27
Low reliability	1	10	34

[1] m = maximum number, for a clutch, of mismatches per excluded trio.

4.4.4 Behaviour-based Mismatches: Estimation of Incorrect Parentage

The first approach that may be used to quantify those mismatches due to laboratory error is to find that value of actual mismatches (i.e. not due to laboratory error) that would be expected to result in the observed number of multiple mismatches. This approach was used by Ashton (1980) where a value of maximum likelihood was found as that producing the minimum $\underline{X}^2$ in a goodness-of-fit test. In this study only 4 cases of exclusion at more than one locus were observed. In a further clutch two sibs were excluded at different loci, but the assumption that all offspring are attributable to a single mating is not being made here, even though there was no direct evidence for multiple mating (through, for example, the detection of three or more maternal alleles). One of the 4 cases included an exclusion at the EST2 locus, which was not included in the calculation of mismatches, and was therefore excluded from the analysis here. Reference to Table 4.17 shows that multiple mismatches totalled only about 25% to 35% of expectation, depending on whether the actual mismatches would be due to single parent or double parent inconsistencies. These give estimates for the proportion of incorrectly excluded individuals among the observed exclusions of 34% and 26% respectively. As these estimates are based on extrapolation from only 3 multiple mismatches they must have a very low degree of reliability. Further, they depend upon expected distributions which, though appearing appropriate on average, showed some deviation from those observed in the simulation study (Section 4.3.4 and Fig. 4.2).

If these estimates were correct, they would lead to estimates for the rate of behaviour-based mismatches of about 5.8% and 5.4%, depending on whether one or both parents are incorrect.

4.4.5 Intraspecific Brood Parasitism

If both parents were incorrect, then the implication would be that a non-parental egg had been 'dumped' in the nest. On the basis of studies involving daily checks of nestboxes, Summers-Smith (1963) and Seel (1968a) concluded that house sparrows lay their eggs on consecutive days, early in the morning. Other workers (e.g. Dawson 1972, Sappington 1975) have not disagreed with this. It was therefore not considered necessary to check boxes daily during egg-laying in this study (see Chapter 2). It is possible that birds in my study populations behaved differently from those studied elsewhere, but the frequency of checking was high enough to make the non-detection of the appearance of more than one egg per day on at least a proportion of those occasions highly unlikely. In fact, on two occasions three eggs were recorded as appearing in two days, but both referred to the start of clutches and these events were interpreted as being due to the non-detection of the first egg laid. The first egg is frequently laid before completion of the nest and may be easily concealed by loose nesting material. Complete families were not obtained in either of these nests.

There are, of course, ways in which egg-dumping might not be accompanied by the appearance of extra eggs. For example, some avian species are known to be indeterminate layers, such that they lay eggs until the clutch attains a particular total size regardless of any losses or additions which occur during the laying period. Others are referred to as determinate layers, since they lay a number of eggs determined by the number of follicles that start to develop in advance of egg laying (see Klomp 1970 for review). If the house sparrow were an indeterminate layer, then some dumped eggs would go unnoticed as the female would lay fewer eggs as a

consequence. Evidence concerning whether house sparrows are determinate or indeterminate layers is conflicting. Witschi (1935) reported egg removal in this species leading to the laying of up to 50 eggs. Schifferli (1976) cited Kreymborg (1911) and Puhlmann (1914) who described deliberate egg removal to induce continuous laying as a method used to kill sparrows, but was himself unable to induce continuous laying in this way, and suggested that the earlier results were more likely to be due to the laying of repeat clutches. Of more interest here is the effect of egg addition; for 10 nests in which Schifferli added 4 eggs to the first laid there was no apparent reduction in the mean total number laid, though some clutch sizes were small. Interestingly, no desertions resulted, which I take to imply that females do not discriminate their own eggs, but irregular laying (extremely unusual in this species) occurred in two instances.

Alternatively, a parasitic female might remove an egg at the time of laying; this is a well-documented behaviour in interspecific parasites (e.g. the European cuckoo, Cuculus canorus, Wyllie 1981) and has been shown to occur frequently (by marking eggs: Evans 1980) during intraspecific parasitism by starlings. Broken eggs were occasionally found on the ground near to sparrow nests during the egg laying period, even though none were known to have disappeared from nearby nests. Removed eggs might in any case be expected to be carried well away from the nest, either by the parasite or else by an appropriate scavenger (such as the starling, commonly nesting close to sparrows). When clutches were deserted, for whatever reason, eggs frequently disappeared without trace. Schifferli (1976) specifically mentions that "no trace of broken eggs could be found in or under the nest-boxes" following desertion due to the

deliberate removal of breeding males.

The size, shape, patterning and colouration of eggs varies markedly, both within and between females (Dawson 1972 and personal observation). The variations observed within clutches are an occasional pronounced shape difference of the first egg of the season (longer and narrower), or pronounced pattern differences, manifested as a paler background colour and larger, less densely distributed spotting, frequently of the last egg of the clutch. This latter phenomenon has been referred to previously (Seel 1968a), and has been frequently recorded in a range of other species (see Yom-Tov 1980). Thus despite differences between females, a subjective assessment of egg differences would have been inadequate to quantify any dumping and only a small sample of clutches was examined closely.

If egg dumping associated with egg removal were to occur commonly (i.e. at a frequency affecting up to 5% of eggs) then it is surprising that this behaviour has not been reported, to my knowledge, by any of the many people who have worked on this species. The only published claims concerning dumping in the house sparrow were made by Manwell and Baker (1975) who recorded two instances of clutches containing eggs which, on the basis of their interpretation of the inheritance of egg albumin proteins, could not have been laid by a single female. Their observation was not, unfortunately, accompanied by ecological data and it remains possible that two females were sharing a nest in each case; such polygynous behaviour within a nest has been reported previously although it is generally rare (Lowther 1979c). On balance, it would appear unlikely that the suggested level of behaviour-based mismatches could be accounted for by intraspecific brood parasitism.

4.4.6 Nonpaternity

If behavioural events resulting in nonpaternity are the sole source of these mismatches, then those mismatches attributable to incorrect female parentage (see Section 4.3.1) will be due to laboratory error. When a particular parent was incorrect, it was more often the male, though the difference did not achieve statistical significance (Section 4.3.1). This suggestion of a difference might be taken as evidence that behaviour-based mismatches result more often from behavioural events involving non-parental males than females. For each laboratory error that excludes a female parent, one excluding a male parent is also expected to occur. Thus for the data presented here (Table 4.7), 8 of 11 cases where attribution of nonpaternity/nonmaternity was made might be expected to be due to laboratory error, suggesting a rate of laboratory error of 73%. This estimate, even if the assumptions on which it is based are accepted, will of course be extremely unreliable because of the small number of available exclusions involving a specific parent. It is, however, close to the estimate of 66% derived from the frequency of multiple mismatches (see above). Thus a frequency of behaviour-based mismatches of about 5%, in this case involving non-parental males only, is again suggested.

Two kinds of event might produce this result: extrapair copulations and unrecognized mate change between egg-laying and the time when parentage was recorded. Such mate change is thought to be unlikely to account for the results for two reasons. First, with regard to the response of a new male taking over a nest already containing a clutch, there is no evidence to suggest that incumbent males are ever ousted, and so their replacement is more likely to be

due to their death (Summers-Smith 1963). In a male removal experiment involving 8 nests, Schifferli (1976) found that most nests failed, but hatching occurred at one where a new male was observed. Second, combining data from both sites, the proportion of males known to remain from previous clutches of the same year as compared with those known to be different was much the same for excluded and non-excluded clutches:

Table 4.20

The number of clutches in which the male parent was known to have remained from a previous clutch at the same nest compared with the number in which it was known to have changed.

	Male same as in previous clutch	**Male different from previous clutch**
Clutches with mismatches	6	3
Clutches without mismatches	22	6

We are left, then, with a suggestion of the occurrence of extra-bond copulations. As pointed out above (Section 4.3.6) few observed copulations in this study took place between positively identified birds. The one instance of copulation of a known male with an (unmarked) non-mate could conceivably have been a polygynous, rather than an extra-bond, copulation. Polygyny has been observed occasionally in this species and is believed to result

from temporary shortages of males and/or nest sites (Summers-Smith 1958, 1963, Lowther 1979c). Of the studies involving marked house sparrows (e.g. North 1968, Sappington 1975, Summers-Smith 1963), only Summers-Smith mentions observing copulations: no extra-bond copulations were seen. Summers-Smith's study was carried out in a suburban area, where the densities of sparrows and their nests during the breeding season was far lower than here. One might predict that the occurrence of extrapair copulations will be density-dependent (either with respect to density of nests or of adult males); a difference between populations would not in that case be unexpected.

It would be difficult to estimate the effect of density from the data presented here; all nest-boxes within each site were probably within the foraging range of all birds at each site. Nest density is difficult to assess as nests other than those in nestboxes would have to be included and as boxes tend to be clumped (around particular buildings) density becomes difficult to quantify. Further, where nestboxes had lower occupancy rates and nest density was obviously lower the numbers of successful nests are too small to allow comparison. There will be temporal as well as spatial density differences however, and at times when nest activity is reduced (and hence active nest density is reduced), the density of adult males not directly involved in nesting will, conversely, increase. The first clutch of the season is well known as the time of the greatest nesting activity in the population and concludes with fledging in June. No difference in the occurrence of mismatches was found between any of the months in which fledging was observed.

A consideration of other ecological aspects was also made in an

attempt to gain some insight into the occurrence of mismatches, whatever their origin might be. For example, it has been shown that male Barbary doves, Streptopelia risoria, can discriminate between those females that have and those that have not been exposed to a different male (Erickson and Zenone 1976, 1978; Zenone, Sims and Erickson 1979; Rissman 1983). This is apparently enabled by behavioural changes in the female following the induction of ovarian steroid secretions by previous exposure to a displaying male. The discriminating ability of males is interpreted as a method of avoiding being cuckolded, but little is known at present of the ability of males to detect infidelity on the part of mated females. An experiment involving the separation of mated pairs of S. risoria for various time periods failed to detect any differences in the behaviour of the male towards his mate whether or not the female had spent the intervening time period with a different male (Rissman 1983). Rissman suggested that the temporarily separated males were more 'cautious' towards their mates following separation, but it is unclear what adaptive value 'caution' might have. If a male's potential paternity is in doubt then one might predict that he should behave in a manner to maximise his personal chances of paternity, either by copulating with the returning mate or else by abandoning the initiated clutch. If house sparrow males have any cuckold detecting ability then their parental care behaviour might well change as a consequence, leading to reduced breeding success. Of course, their behaviour in such instances might be to desert or drive away the female, in which case, even though a female may raise a brood alone, they will not form part of the sample considered here. A comparison of breeding success between clutches having and not having mismatches is confounded by the effect increased breeding

success will have upon increasing the probability of discovering a mismatch.

4.5 Conclusions

Although many mismatches were observed, a careful examination of the data has revealed a number of problems which confound their interpretation. The most parsimonious interpretation is that nonpaternity occurred amongst nestlings at the two study sites at a rate of about 6%. This may not be an excessive rate for a monogamous species: a higher rate in a mammal species regarded as monogamous has been described as surprisingly low (Foltz 1981a). The finding that calculated multiple-locus exclusion probabilities over-estimated the rate of exclusion in simulations of nonparentage, which used the genotypes of sibships that were actually observed, was one of the unexpected problems. The actual probability of exclusion depends upon the distribution of multiple locus genotypes, which appeared to be nonrandom in the two study populations, though there was very little evidence for digenic disequilibria in these populations (Chapter 5). The use of simulations of the type introduced here to test the applicability of estimated exclusion probabilities would appear to be a desirable feature of this kind of study. The testing of the distribution of the number of mismatches per individual with expectation would appear to be essential: in the few studies, including this one, where this has been done this has led to the discovery of errors. If the possibility of errors is ignored then estimates of nonparentage are liable to be seriously flawed.

4.6 Summary

357 nestlings were sampled for 7 polymorphic protein loci in 124 clutches where both putative parents were also sampled. 12.9% of the nestling genotypes were interpreted as being genetically incompatible with those of their parents. Exclusion probabilities were calculated as 43-51% for nonpaternity and 59-67% for nonparentage. The applicability of these estimated probabilities was tested by the random reassortment and comparison of observed parental genotypes among observed sibship genotypes. Significantly fewer nestlings were excluded in these simulations than expected from calculated exclusion probabilities, though the distribution of multiple mismatches among the detected simulated cases of nonpaternity and nonparentage did not differ from expectation. The distribution of multiple mismatches did differ from expectation, however, in the actual mismatch data, implying the occurrence of errors in the attribution of mismatches. Possible sources of error, which must have included some form of laboratory error, were considered and discussed. The most parsimonious interpretation of those mismatches that did not appear to be due to errors was that they resulted from a rate of nonpaternity of about 6%. No heterogeneity in the rate of mismatches was observed within or among breeding seasons or sites.

CHAPTER 5

POPULATION GENETICS

5.1 Introduction

The advent of electrophoretic techniques has allowed the quantification, in a variety of species, of the genetic relationships within and among populations (see Weir and Cockerham 1984 for a list of examples). It has been argued that the observed levels of variability and differentiation are not at variance with the predictions of neutral theory (Kimura 1983). Three studies of the genetic structure of house sparrow populations have reached a similar conclusion. Each of these studies involved the comparison of electrophoretically-determined estimates of genetic differentiation with estimates predicted by neutral theory from available ecological and historical data concerning the species (Fleischer 1983, Parkin and Cole 1984a, 1984b; refer to Sections 1.2 and 1.4). Although it was not found to be necessary to invoke the operation of factors other than gene-flow, random genetic drift and randomness of mating to explain the genetic differentiation observed among house sparrow populations, the possible importance of other factors was not excluded. Other potentially important forces include natural selection, and interactions among loci.

Natural selection has been implicated in the maintenance of several protein polymorphisms (see Section 1.2). Some of this evidence has been obtained by comparing allele frequencies in different populations relative to components of the environment, and finding that allele frequencies in an independent set of populations may be partly predicted from similar environmental data. In other cases evidence for selection has been obtained by the detailed

ecological genetic study of individuals within populations. Studies of this kind have for practical reasons largely concentrated on populations of small mammals and birds (Section 1.2), and the house sparrow was believed to be a particularly suitable species for this kind of investigation (see Section 1.4).

The methods for the non-destructive sampling of 7 protein loci have been described (Chapter 2), and the inheritance of these loci investigated (Chapter 3). Distorted segregation ratios among fledglings suggested the operation of selection against the rarer alleles at the PEPD3 and GP1 loci. This selection could be either gametic or zygotic; in either instance its effects upon allele frequency prior to fledging would be directional rather than balanced. The maintenance of these polymorphisms implies either the operation of some other form of selection or else differential gene-flow at a later stage in the life cycle. A comparison of genic distributions among different age and sex classes potentially allows the detection of these nonrandom effects.

Selection may also vary temporally (see Ford 1975 for examples). Fecundity in the house sparrow is likely to vary in response to differing environmental conditions. For example, Anderson (1977) found that fledging success in a North American population of house sparrows increased during a period of food superabundance in an emergence year of periodical cicadas. The strength and nature of selective forces might be expected to similarly vary from year to year, and consequently changes in genic distributions might result. Temporal variation in allele frequencies at enzyme loci has been reported in at least one avian species: the eared dove, Zenaida auriculata, in South America (de Caminos et al. 1981).

It has been argued that comparisons of genic distributions should be made among genotypes rather than alleles as most forms of selection are expected to act upon genotypes (De Benedictus 1978). Where selection or gene-flow operates with respect to more than one genotype, however, statistical tests may be more likely to detect their effects if allelic distributions are considered. Both genotypes and alleles were examined, where possible, in the data for the two populations of house sparrows presented here; in general, results are presented for comparisons of phenotypes (electromorphs) only, but are also presented for alleles if any differences were found between the comparisons made in each way. With the exception of some phenotypic classes for EST2, the phenotypes were interpretable as genotypes (Chapter 3).

5.2 Methods

Detailed descriptions of the field and laboratory methods (Chapter 2) and of the seven protein polymorphisms (Chapter 3) have been provided above. The principal statistical methods used were those of goodness of fit tests and tests of homogeneity. Goodness of fit tests were carried out using the log likelihood ratio ($\underline{G}$) test (Sokal and Rohlf 1981). The values obtained for $\underline{G}_d$ are distributed approximately as the χ^2_d distribution for the appropriate degrees of freedom, $\underline{d}$. In tests of correspondence of genotypic proportions with Hardy Weinberg expectations, the degrees of freedom were found as the number of possible genotypes less the number of observed alleles (Cavalli-Sforza and Bodmer 1971), i.e. $\underline{d} = (\underline{a}^2 - \underline{a})/2$ where $\underline{a}$ is the number of alleles at the tested locus. Tests of homogeneity were carried out using the familiar χ^2 test. The sample statistic obtained, $\underline{X}^2_d$, is again distributed

approximately as the appropriate χ^2_d distribution (Sokal and Rohlf 1981).

Tests of homogeneity and goodness of fit are both subject to distortion when the expected values are low. Numerical simulations of G and χ^2 tests by Conahan (1970, cited by Sokal and Rohlf 1981) suggest that ideally no expected frequency in a goodness of fit test should be less than 10, but for practical purposes a minimum of 5 is adequate. A minimum expected value of 3 was sufficient where the number of tested classes was 5 or more; in these circumstances the G test performed better than the χ^2 test. Sokal and Rohlf (1981) do not even discuss minimum values for expecteds in G tests of homogeneity, but a minimum value of 1 is believed to be conservative in χ^2 tests (Everitt 1977 p.40). Thus the magnitude of expected values that will produce serious distortion would appear to be greater in goodness of fit than heterogeneity tests, an effect not emphasised in popular statistical texts. These recommendations of minimum expecteds for goodness of fit and homogeneity tests were therefore followed throughout this study.

Most contingency tables were constructed using the CROSSTABS program (SPSS: Nie et al. 1975) This program was also used to obtain initial estimates of χ^2 in tests of homogeneity. Where values of X^2 were low ($X^2 < 3.84$), there could be no significant heterogeneity regardless of the magnitude of expected values. Where larger values of X^2 were obtained ($X^2 > 3.84$) and some expected values were below 1, if the difference ($X^2 - 3.84$) was more than could be accounted for by the cells containing low expecteds, then X^2 was recalculated omitting rows containing those cells. Similarly, if the value for G was low ($G < 3.84$) in goodness of fit tests, then no pooling or other procedures were followed even if

some expecteds were small.

The statistical analysis of allozyme data included a series of χ^2 contingency tests for homogeneity among the samples. These were conducted initially without any correction for sib effects as although non-independence is expected to result in increased heterogeneity, it was also expected that in many such instances the values obtained for X^2 would remain small. In this way the initial computation was minimised. There was no a priori expectation with respect to possible interaction among the variables and they were therefore analysed separately.

5.3 Results

5.3.1 The Data

The electrophoretic data are presented for each locus with respect to site, year, age and sex categories (Appendix 2). The group totals in some instances differ slightly among loci due to the occasional depletion of smaller tissue samples. The substantial reduction in numbers for PEPD3 results from an inability to score all PEPD3 alleles unambiguously in $PEPD2^{A/B}$ individuals, and these PEPD2 heterozygotes have therefore been excluded from the data for PEPD3 (see Section 3.1.3). The sexes of nestlings were not known. Those of juveniles were known only occasionally and then more often for males as the plumage characteristics of adult males represent an easily recognisable change from the juvenile plumage type (see Section 2.1.4). No attempt has, therefore, been made to analyse the sexes separately in nestlings or juveniles.

The nestlings were almost always part of a completely-sampled sibship, and the electromorph frequencies shown do not therefore

represent entirely independent data. Non-independence due to familial relationships will be discussed further at appropriate parts of the analysis. Many of the parents of these siblings are included in the adult samples. Retrapped (or re-observed) individuals were not included in the analysis; thus none of the juveniles were known to be related to any of the adults, but on very rare occasions, were known to be part of the breeding sample in a subsequent year and have offspring included in the appropriate pullus groups (3 instances).

To summarise, the adult and juvenile groups are believed to represent essentially random samples, the nestlings are those fledging from nests in the study nestboxes and include many known (and probably many unknown) sibs, the juveniles are only rarely known to be related to the nestlings, and the adults often are.

5.3.2 Allele Frequencies

Allele frequencies were initially estimated for all loci by simple gene counting, on the hypothesis that alleles were codominant and each electromorph was interpretable as a single genotype. At the EST2 locus there was an apparent, extremely significant deviation from Hardy Weinberg ratios in the larger of the two adult samples (SB: $\underline{N} = 351$, $\underline{G}_3 = 21.22$, $\underline{p} < 0.001$), and this was attributable to an excess of homozygotes ($\underline{G}_1 = 10.26$, $\underline{p} < 0.005$). This pointed to the presence of a concealed null allele (see Section 3.1.3). Further evidence for null alleles at EST2 was obtained from family studies (Section 3.2.1). Allele frequencies have therefore been estimated for EST2 by an iterative maximum-likelihood method (Li 1955) using a computer program written by Dr. J. Rostron (North East London Polytechnic). The method estimates directly the

frequencies p, q, r of the non-null types EST^A, EST^B, EST^C respectively, together with their standard errors. The frequency s of the null allele $EST2^O$ is found by subtraction as $1-p-q-r$.

The allele frequencies for each locus are also presented with the electromorph data (Appendix 2). The standard errors for the EST2 non-null allele frequencies were found as $\sqrt{\{1 - (1 - p')^2\}/4N}$ where p' was the allele frequency and N was the sample size (Elandt-Johnson 1971 p.397). The standard errors for the null allele frequencies may be found easily as $\sqrt{(1 - s^2)/4N}$ (Elandt-Johnson *op. cit.*), and those for all alleles at the other loci as $\sqrt{p'(1 - p')/2N}$ (Falconer 1981). All the standard errors calculated using these formulae will be underestimates in the case of nestling samples due to the intuitively obvious correlation among sibs (see below). Though confidence limits will be affected by relatedness, the estimated allele frequencies for sibs should represent maximum-likelihood values provided that the sampled sibships are a random sample of all the sibships in the populations (Cotterman 1954). The maximum-likelihood estimates for the EST2 allele frequencies were tested for agreement with Hardy Weinberg expectations (Appendix 2). One test produced a significant G statistic (SB adult males: $G_3 = 9.16$, $p < 0.05$), but this cannot be regarded as truly significant as it was due to the occurrence of some low expected values.

Family data concerning the other six loci was in agreement with the codominant model applied here (Section 3.2.1). In general, there was also close agreement with Hardy Weinberg ratios for these loci (Appendix 2). Only four G values exceeded the 5% significance level for the appropriate degrees of freedom, and two of these were attributable to very low expected frequencies (< 1). It is

interesting that both the remaining two deviations applied to IDHC for 1982 nestlings, and both reflected an apparent deficiency of heterozygotes. The magnitude of the departure will, however, be at least partly due to non-independence of sampling among nestlings, particularly marked at this locus (see below).

5.3.3 Adults: Sex and Year Classes

The sexes were distinguishable only in adults and were therefore analysed first. Only one out of 42 comparisons of electromorph frequencies between the sexes was significant at the 5% probability level (Table 5.1). It was concluded that there was no heterogeneity between sexes sampled within year and site classes at any locus.

The ages of the adult house sparrows could not be precisely determined (see Section 2.1.4). With respect to adults, the biological significance of year classes as used here (the year of first capture) is expected to be limited. More adults when first captured are expected to have originated in the previous year than in any other year, but most are likely to be aged two years or older (from data in Summers-Smith 1963). As the study progressed some adults were known to have survived into at least their third or fourth year, but as few birds sampled in their first year were reencountered during the limited number of subsequent breeding seasons available, very few adults were of known age and a comparison among adult age classes was not considered to be worthwhile. The sample years were, however, tested for heterogeneity within sites.

Among the 14 tests one was significant at the 1% level (Table 5.2). This was the result of an anomalous excess of $\underline{GPl^{C}}$ alleles at

Table 5.1

Comparison of electromorph frequencies in sexes of adults within year and site classes. d = degrees of freedom.

Site	Year	Locus													
		6PGD		PEPD3		PEPD2		IDHC		PEPT		GPl		EST2	
		X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d
SB	1980	1.20	2	0.21	3	3.54	3	3.22	2	1.33	3	2.51	3	2.01	3
	1981	0.15	1	3.39	3	1.83	2	1.73	2	4.92	4	3.44	5	4.00	6
	1982	0.16	1	0.69	1	1.81	1	1.04	2	8.62	4	4.93	5	2.02	3
BR	1980	2.09	1	0.26	2	1.69	1	2.15	2	4.17	4	0.68	3	1.75	4
	1981	0.16	1	1.41	3	0.00	1	2.93	2	3.22	4	2.09	5	2.00	3
	1982	0.00	1	0.38	2	0.15	1	6.31*	2	4.09	4	0.97	3	4.83	3

Table 5.2

Comparison of electromorph frequencies of adults at each site among year classes. d = degrees of freedom.

Site	Locus													
	6PDGD		PEPD3		PEPD2		IDHC		PEPT		GPl		EST2	
	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d
SB	4.19	4	7.10	6	5.82	6	6.13	4	6.21	10	6.63	10	9.28	4
BR	2.28	2	5.06	6	1.18	2	6.54	4	6.75	10	13.49*	6	7.39	10

* $p < 0.05$

Brackenhurst in 1981. The numbers (Table 5.3) were, however, rather small in this instance and a general conclusion of homogeneity was reached. Adults were therefore pooled across years prior to further analyses, and the pooled genotypes, for the 6 codominant loci, were tested for agreement with Hardy Weinberg expectations (Table 5.4). The observed phenotype frequencies for the single locus exhibiting dominance, EST2, did not depart significantly from those expected from the maximum-likelihood allele frequencies (SB: G_3 = 4.07, n.s.; BR: G_3 = 0.019, n.s.). Curiously, the only deviation ($p < 0.05$) was again associated with those genotypes including the GPI^C allele, but in this case at Sutton Bonington. An evaluation of this result is compounded by the low expected frequencies of these genotypes.

The lack of departure from Hardy Weinberg equilibrium does not necessarily exclude the possibility of a significant degree of inbreeding. The inbreeding coefficient F for a locus may be estimated as

$$\hat{F} = 1 - (H_O/H_E)$$

where H_O is the observed number of heterozygotes and H_E is the expected number (Crow and Kimura 1970). Positive values for $\hat{F}$ result from a deficit of heterozygotes, and negative values from an excess. H_E is determined as $(1 - \sum p_i^2)N$ where p_i is the frequency of the ith allele at a locus and N is the sample size. A less biased estimator of H_E, H_E', has been employed here and was found as $H_E' = H_E\{2N/(2N - 1)\}$ (Levene 1949, Crow and Kimura 1970).

The values for $\hat{F}$ for adults are presented in Table 5.5. Where the value for G for the Hardy Weinberg test was low (i.e. $G < 3.84$), it was known that the estimate $\hat{F}$ would not be significant. Observed and expected heterozygosities were compared in the one instance where a significant departure from Hardy Weinberg was observed - GPI

Table 5.3

Comparison of GP1 genotype and allele frequencies among years for Brackenhurst adults.

Year	Genotypes						Alleles		
	AA	AB	AC	BB	BC	CC	A	B	C
1980	6	25	0	70	2	0	35	165	2
1981	4	15	1	39	8	1	24	107	11
1982	2	25	0	42	2	0	29	111	2
Hetero-geneity test	$X^{2*}_{6} = 13.5$ $p = 0.036$						$X^{2}_{4} = 16.1$ $p < 0.005$		

* test excludes AC and CC classes

Table 5.4

Results of goodness of fit G tests for agreement of adult genotype frequencies with Hardy Weinberg proportions.
O = observed, E = expected, df = degrees of freedom.

Locus	Site		Genotypes														G_{HW}	df
			AA	AB	AC	AD	AE	BB	BC	BD	BE	CC	CD	CE	DD	DE	EE		
6PGD	SB	O	1	16				340											
		E	0.2	17.6				339.2										1.56	1
	BR	O	0	16				227											
		E	0.3	15.5				227.3										0.54	1
PEPD3	SB	O	0	4		0		278		34					3				
		E	0.0	3.7		0.3		276.5		37.2					1.3			2.59	3
	BR	O	0	7	0	0		190	1	23		0	0		0				
		E	0.1	6.5	0.0	0.4		191.1	0.9	21.4		0.0	0.1		0.6			2.34	6

Table 5.4. (cont'd)

Locus	Site		Genotypes														G_{HW}	df
			AA	AB	AC	AD	AE	BB	BC	BD	BE	CC	CD	CE	DD	DE	EE		
PEPD2	SB	O	1	37	0			315	4			0						0.48	3
		E	1.1	36.7	0.2			315.3	3.8			0.0							
	BR	O	0	22				221										1.04	1
		E	0.5	21.0				221.5											
IDHC	SB	O	194	131				30										1.32	1
		E	189.7	139.6				25.7											
	BR	O	117	102				24										0.07	1
		E	116.2	103.7				23.2											

Table 5.4 (con'td)

Locus	Site		Genotypes														G_{HW}	df
			AA	AB	AC	AD	AE	BB	BC	BD	BE	CC	CD	CE	DD	DE	EE		
PEPT	SB	O	0	0	0	2	0	0	1	20	0	0	18	0	313	3	0		
		E	0.0	0.1	0.1	1.9	0.0	0.3	0.6	19.7	0.1	0.3	17.8	0.1	313.4	2.8	0.0	2.03	10
	BR	O						2	1	16	0	0	14	0	200	10	0		
		E						0.5	0.7	19.0	0.4	0.2	13.6	0.3	199.2	9.1	0.1	5.73	6
GP1	SB	O	11	86	7			240	12			2							
		E	9.2	92.8	3.7			233.3	18.6			0.4						9.51*	3
	BR	O	12	63	1			151	12			1							
		E	8.1	69.1	2.8			148.1	11.8			0.2						5.13	3

* $p < 0.05$

Table 5.5

Observed (H_O) and expected (H_E') numbers of heterozygotes and estimates of inbreeding coefficients in each study population for 6 codominant loci.

Locus	Site					
	SB			BR		
	H_O	H_E'	$\hat{F}$	H_O	H_E'	$\hat{F}$
6PGD	16	17.6	0.0922	16	15.56	-0.0301
PEPD3	38	40.9	0.0724	31	28.87	-0.0740
PEPD2	41	40.7	-0.0060	22	21.04	-0.0452
IDHC	131	139.6	0.0629	102	103.91	0.0184
PEPT	44	43.2	-0.0171	41	43.19	0.0507
GPI	105	115.1	0.0865	76	83.87	0.0939
Mean			0.0485			0.0023

at Sutton Bonington - but no significant inbreeding was found ($\underline{G}_1$ = 1.33). Thus the single departure from Hardy Weinberg expectation was an effect associated with the $\underline{GPI^C}$ allele directly and not due to any inbreeding or other factor affecting heterozygosity. Evidence for possible selection at GPI has been provided earlier (Section 3.2.1) and this will be discussed further below.

5.3.4 First-year Birds: Year Classes

A comparison of year classes for nestlings and juveniles is potentially more valuable than the same analysis (as just conducted) for adults. As already stated, many of the nestlings will be members of sibships, and tests for heterogeneity between groups of nestlings ideally should allow for this wherever possible. The effect of any relatedness will be to increase the variance beyond that expected if sampling were independent. However, it is extremely unlikely that true heterogeneity among groups will be counteracted by the effects of non-allowance for non-independence. To minimise computation, therefore, comparisons involving nestlings were first made by treating nestlings as though they were independent.

The results for heterogeneity χ^2 tests among year classes and within sites for the electromorphs of nestlings and juveniles separately are presented in Table 5.6. In several instances where the loci are less heterozygous, the number of juveniles sampled within a year was too small to allow the test. Significant heterogeneity ($\underline{p} < 0.01$) was suggested for IDHC nestlings at both sites, and slight heterogeneity ($\underline{p} < 0.025$) for EST2 in Sutton Bonington nestlings. A comparison of allele frequencies produced similarly significant results for IDHC (SB nestlings: $\underline{X}^2_2 = 10.36$, $\underline{p}$

Table 5.6

Comparison of electromorph frequencies of (i) nestlings and (ii) juveniles at each site among year classes. d = degrees of freedom.

Site	Sample	Locus													
		6PGD		PEPD3		PEPD2		IDHC		PEPT		GPl		EST2	
		X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d
SB	(i)	6.58	4	7.70	6	3.84	4	14.92**	4	10.73	8	13.43	8	7.77*	2
	(ii)	-		-		-		2.95	4	-		0.16	2	-	
BR	(i)	3.36	4	10.98	8	0.63	2	13.78**	4	6.21	10	11.24	8	8.26	4
	(ii)	-		-		-		6.71	4	-		6.37	4	-	

* $p < 0.05$ ** $p < 0.01$

= 0.006; BR nestlings: $\underline{X}^2_2 = 8.73$, $\underline{p} = 0.013$). The largest degree of heterogeneity among juvenile allele frequencies was for IDHC at Brackenhurst ($\underline{X}^2_2 = 5.25$, $\underline{p} = 0.073$), but the differences among years did not parallel those for nestlings at the same site. When Brackenhurst nestling and juvenile IDHC alleles were pooled within years (as all known birds of the year) the heterogeneity was not significant ($\underline{X}^2_2 = 5.04$, $\underline{p} = 0.080$). Heterogeneity of nestling allele frequencies among years at Sutton Bonington was also suggested for 6PGD ($\underline{X}^2_2 = 6.01$, $\underline{p} = 0.050$) and GP1 ($\underline{X}^2_4 = 9.06$, $\underline{p} = 0.060$). The IDHC results are of most potential interest as they produce the largest $\underline{X}^2$ values at each site. The suggested fluctuations in allele frequency (Appendix 1) are not, however, consistent between the sites. These results will be investigated more rigorously below.

One way of increasing independence would be to randomly select a single individual from each sibship, but this would be very wasteful of information. Cotterman (1947, 1954) has discussed the estimation of frequencies and their variances for codominant allele data obtained from families. In principle, the alleles observed amongst a cohort of full sibs may not represent more than a theoretical maximum of four independent sampling events (the 4 original parental alleles). The number of independent events that a sibship represents increases asymptotically with the size of the sibship. Thus a weighting can be applied to each observed allele in a sibship, and the larger sibships will not then have an undue effect upon the allele frequency estimates. The estimates for gene frequencies are not expected to be affected by the weighting procedure if sibships are sampled randomly. Variance estimates will, however, increase with an increase in mean sibship size for a

constant sample size. Similarly, weighted totals will decrease with an increase in sibship size.

Cotterman (1947) derived the weighting, w_o for each sib allele in a sibship of size s as $2/(s + 1)$. The total weight W_o appropriate to a sibship of size s may be found as the sum of the weights for each allele, or as $2sw_o$. When frequencies and their precision estimates for the entire population are required, different weightings become applicable when only one parent is known, and offspring are ignored if both parents are known. For clutches sampled here one or both parents are frequently known, and clutches within and among years are often known to have the same parent. For the analysis of nestlings among years only the nestling allele frequencies are of immediate interest, however, and the weightings appropriate to the case of no known parents will be applied. Most clutches at the same nest within a breeding season are expected to be the progeny of the same pair (Summers-Smith 1963; personal observation). Nestlings at the same nest within a year were therefore treated as full sibs unless a different pair was known to have taken the nest over. Such complete replacement of pairs only occurred three times. Relationships among sibships in successive years do not affect this analysis.

The analysis for IDHC at each site and for each year is set out in Table 5.7. The weighted totals were tested for heterogeneity of allele frequencies among years within each site and none was found (SB: $X^2_2 = 0.65$, $p = 0.72$; BR: $X^2_2 = 4.15$, $p = 0.126$). Thus no heterogeneity actually existed, and that estimated when lack of independence was ignored (even achieving significance at the 1% level) was due to that lack of independence.

The total weightings were close to 40% of the original allele

Table 5.7

Estimation of weighted allele frequencies for IDHC in nestlings, correcting for correlation among sibs. s = sibship size, w = weighting, N_s = number of clutches of size s, N_A = number of A alleles and N_B = number of B alleles. $2N_sw$ = weighted total. Refer to text for further explanation.

Year	s	w	Site											
			BR						SB					
			N_s	Alleles N_A	Alleles N_B	Weighted wN_A	Alleles wN_B	$2N_sw$	Ns	Alleles N_A	Alleles N_B	Weighted WN_A	Alleles wN_B	$2N_sw$
1980	1	1.000	1	2	0	2.00	0.00	2.00	-					
	2	0.667	1	4	0	2.67	0.00	2.67	2	5	3	3.33	2.00	5.33
	3	0.500	1	5	1	2.50	0.50	3.00	2	9	3	4.50	1.50	6.00
	4	0.400	2	13	3	5.20	1.20	6.40	3	15	9	6.00	3.60	9.60
	5	0.333	1	5	5	1.67	1.67	2.33	1	6	4	2.00	1.33	3.33
	6	0.286	-						-					
	7	0.250	-						2	17	11	4.25	2.75	7.00
	8	0.222	-						2	32	0	7.11	0.00	7.11
	Total					14.03	3.37	17.40				27.19	11.18	38.38

Table 5.7. contd.

Year	s	w	Site											
			BR						SB					
				Alleles		Weighted	Alleles			Alleles		Weighted	Alleles	
			N_s	N_A	N_B	wN_A	wN_B	$2N_sw$	Ns	N_A	N_B	WN_A	wN_B	$2N_sw$
1981	1	1.000	2	3	1	3.00	1.00	4.00	7	12	2	12.00	2.00	14.00
	2	0.667	7	24	4	16.00	2.33	18.67	10	22	18	14.67	12.00	26.67
	3	0.500	9	34	20	17.00	10.00	27.00	11	52	14	26.00	7.00	33.00
	4	0.400	9	56	16	22.40	6.40	28.80	8	56	8	22.40	32.00	25.60
	5	0.333	1	9	1	3.00	0.33	3.33	7	53	17	17.67	5.67	23.33
	6	0.286	3	24	12	6.86	3.43	10.29	2	22	2	6.29	0.57	6.86
	7	0.250	1	9	5	2.25	1.25	3.50	1	12	2	3.00	0.50	3.50
	8	0.222	-						1	12	4	2.67	0.89	3.56
	11	0.167	1	17	5	2.83	0.83	3.67	1	20	2	3.33	0.33	3.67
		Total				73.34	25.91	99.25				108.02	32.16	140.18

Table 5.7 contd.

Year	s	w	Site											
			BR						SB					
				Alleles		Weighted	Alleles			Alleles		Weighted	Alleles	
			N_s	N_A	N_B	wN_A	wN_B	$2N_sw$	Ns	N_A	N_B	WN_A	wN_B	$2N_sw$
1982	1	1.000	1	0	2	0.00	2.00	2.00	5	9	1	9.00	1.00	10.00
	2	0.667	13	29	23	19.33	15.33	34.67	15	42	18	28.00	12.00	40.00
	3	0.500	8	32	16	16.00	8.00	24.00	10	48	12	24.00	6.00	30.00
	4	0.400	10	55	25	22.00	10.00	32.00	10	62	18	24.80	7.20	32.00
	5	0.333	2	7	13	2.33	4.33	6.67	5	39	11	13.00	3.67	16.67
	6	0.286	5	54	6	15.90	1.71	17.14	3	18	18	5.14	5.14	10.29
	7	0.250	1	8	6	2.00	1.50	3.50	6	73	11	18.25	2.75	21.00
	8	0.222	4	36	28	8.00	6.22	14.22	2	20	12	4.44	2.67	7.11
	10	0.182	-						1	7	13	1.27	2.36	3.64
		Total				85.09	49.10	134.19				127.91	42.79	170.70

totals (from Appendix 2). If the unweighted frequency estimates within the samples are taken as correct, and if all cells in the contingency table are weighted equally, the χ^2 estimate ($\underline{X}^2$) will be reduced by about 60%. In these circumstances values for χ^2 for unweighted data will need to be significant at, or near to, the 0.1% level for the result to remain significant when the data are correctly weighted. The requirement that all cells be equally affected is thought to be conservative: a greater reduction in $\underline{X}^2$ might in fact be expected as those cells producing the biggest $\underline{X}^2$ are likely to be those where lack of independence is most distorting. It is therefore felt to be unnecessary to carry out the weighted analysis for the other loci where heterogeneity among years was initially suggested. It may be safely concluded that there was no significant heterogeneity among years for any age class at any locus.

5.3.5 Comparisons Among Age Classes

It has just been shown that on the basis of three year's data there is no evidence for changes in gene frequency from year to year. The possibility of consistent changes within generations has still to be considered.

As before, the data were initially analysed without allowance for non-independence. The analysis was conducted twice, both with and without the inclusion of the juvenile class, as the totals for juveniles were often small. The results relating to electromorphs are presented in Table 5.8. It can be seen that the strongest suggestion of a difference was for GPI, and that this applied to the samples at both sites. The differences in GPI allele frequency between adults and nestlings were, however, opposite in nature at

Table 5.8

Comparison of electromorph frequencies among age classes within sites. (i) adult and nestling classes only (ii) adult, juvenile and nestling classes. d = degrees of freedom.

Site	Comparison	Locus													
		6PGD		PEPD3		PEPD2		IDHC		PEPT		GPl		EST2	
		X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d
SB	(i)	0.36	2	3.15	3	1.72	3	0.38	2	4.19	5	9.09	5	8.85†	5
	(ii)	0.43	4	5.31	6	7.31	6	1.77	4	5.88	10	7.27*†	1	12.98	12
BR	(i)	0.92	2	2.44	4	0.39	1	1.11	2	8.41	7	13.21**†	4	4.31	5
	(ii)	1.35	4	4.66	8	2.27	2	5.37	4	5.84†	6	15.83*†	6	7.20	4

* $p < 0.05$ ** $p < 0.01$

† excludes electromorph classes with low expecteds (< 1)

the two sites. Weightings for alleles within years have been estimated for GP1 in nestlings, and when the weighted values for each year were compared in turn with the overall adult estimates for each site there were no significant frequency differences between adults and nestlings (Table 5.9). Similarly, no difference was found when the nestling weightings were summed across years (SB: $\underline{X}^2_2 = 3.72$, $\underline{p} = 0.168$; BR: $\underline{X}^2_2 = 0.60$, $\underline{p} = 0.741$). This is believed to be a conservative test as many parents will have bred in successive years and some lack of independence will therefore not have been allowed for. It was concluded that there was no heterogeneity of gene frequencies among age classes.

5.3.6 Site Comparisons

It has been shown (above) that there was no significant heterogeneity among year or age classes within either site. There was, however, known to be a high degree of adult-nestling correlation in the samples. The data have therefore been combined across years, and adults and juveniles have been combined but will be analysed independently of nestlings. This initial analysis, as before, treated nestlings as independent. The results for electromorph comparisons are presented in Table 5.10 and for comparisons of observed heterozygosity in Table 5.11.

A significant difference between genotype frequencies in the adult samples was found for PEPT. This difference was also apparent when alleles were considered ($\underline{X}^2_3 = 9.22$, $\underline{p} = 0.027$), and moreso when adults and juveniles were combined ($\underline{X}^2_4 = 15.58$, $\underline{p} = 0.004$). This was paralleled by a significant difference in heterozygosity for PEPT in the combined adult/juvenile samples ($\underline{X}^2_1 = 4.08$, $\underline{p} = 0.043$). PEPD2 genotype frequencies differed significantly in

Table 5.9

Adult allele frequencies for GP1 at each site (all years combined), compared with weighted scores for nestling frequencies in each year (refer to text and Table 5.7 for details). N = nestlings

Site	Class	Year	Alleles			X^2_2	p
			A	B	C		
BR	Adults	All	88	377	15	-	-
	N	1980	4.27	11.83	1.30	5.79	0.055
	N	1981	25.89	68.06	5.50	0.53	0.767
	N	1982	34.84	96.51	2.85	2.68	0.262
SB	Adults	All	115	578	23	-	-
	N	1980	2.61	33.58	2.18	1.26	0.532
	N	1981	19.66	117.87	1.70	1.96	0.375
	N	1982	20.26	146.19	4.25	0.19	0.909

Table 5.10

Comparison of observed electromorph frequencies in each study population for different age classes. d = degrees of freedom.

Age class	Locus													
	6PGD		PEPD3		PEPD2		IDHC		PEPT		GP1		EST2	
	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d	X^2	d
Adult	1.92	2	5.90	4	3.77	3	2.45	2	10.96*†	3	5.62	5	2.78	6
Juvenile	0.00	1	2.14	2	6.21*	1	4.59	2	4.54	5	5.78	4	0.75	2
Nestling	4.08	3	8.29	4	2.14	2	8.03*	2	12.22*‡	4	50.05***	4	11.11*‡	3

* $p < 0.05$ *** $p < 0.001$

† excludes 6 individuals in electromorph classes containing low expecteds.

‡ excludes 3 individuals in electromorph classes containing low expecteds.

ǂ excludes 4 individuals in electromorph classes containing low expecteds.

Table 5.11

Comparison of observed heterozygosities per locus between study populations for each age class.
N = minimum sample size.

Locus	Age class								
	Nestlings			Juveniles			Adults		
	SB	BR	X^2_1	SB	BR	X^2_1	SB	BR	X^2_1
6PGD	0.037	0.061	2.50	0.043	0.047	0.00	0.045	0.066	1.27
PEPD3	0.092	0.130	2.46	0.000	0.115	2.14	0.119	0.140	0.52
PEPD2	0.116	0.106	0.16	0.261	0.047	8.35**	0.115	0.091	0.91
IDHC	0.360	0.395	0.94	0.273	0.453	2.21	0.369	0.420	1.56
PEPT	0.095	0.139	3.39	0.043	0.187	2.76	0.123	0.169	2.46
GP1	0.233	0.429	32.42****	0.400	0.359	0.11	0.293	0.317	0.37
N	392	277		17	61		319	221	

** $p < 0.01$ **** $p < 0.0001$

juveniles (Table 5.10) and the differences approached significance for the combined adult/juvenile samples for PEPD2 ($\underline{X}^2_2 = 5.34$, $\underline{p} = 0.069$) and GP1 ($\underline{X}^2_5 = 10.28$, $\underline{p} = 0.068$). These results agreed with those for comparisons of allele frequencies (PEPD2: $\underline{X}^2_2 = 5.77$, $\underline{p} = 0.056$; GP1: $\underline{X}^2_2 = 5.34$, $\underline{p} = 0.069$) but only with those for comparisons of heterozygosity in the case of PEPD2 ($\underline{X}^2_1 = 3.23$, $\underline{p} = 0.072$). There was no suggestion of a difference in heterozygosity for GP1 ($\underline{X}^2_1 = 0.56$, $\underline{p} = 0.454$) although heterozygosity was expected to be higher at Brackenhurst; this reflected a larger, but non-significant, deficiency of heterozygotes at Brackenhurst than at Sutton Bonington (Section 5.3.3: Tables 5.4 and 5.5).

Though the difference in heterozygosity per locus achieved statistical significance for PEPT alone, five of the six estimates were higher for the Brackenhurst sample. (The seventh locus, EST2, has not been included here as many null heterozygotes were not directly observable. Frequencies of the detectable EST2 heterozygotes were approximately equal: $\underline{X}^2_1 = 0.00$, $\underline{p} > 0.95$.) Combining probabilities (Fisher's method: Sokal and Rohlf 1981) suggests an overall difference in heterozygosity between the populations ($\underline{p} < 0.05$). The combined probabilities for differences between allele frequencies for 6 loci and between electromorph frequencies for 7 loci (including EST2) were both significant ($\underline{p} < 0.005$ and $\underline{p} < 0.025$ respectively).

Further evidence for heterozygosity differences was obtained from a consideration of all age classes together by examining the number of cases, for each age class, in which heterozygosity per locus was higher at each site (from Table 5.11). Observed heterozygosity at Brackenhurst was higher in 14 of 18 comparisons (2-tailed binomial probability = 0.031). Expected (or 'calculated')

heterozygosities were similarly higher at Brackenhurst for 14 cases out of 18, though not agreeing with the comparisons of observed heterozygosity in every instance. That the comparisons using observed and expected heterozygosities should be in overall agreement was to be expected from the general lack of departure from Hardy Weinberg equilibrium in these samples.

The juveniles and adult male and female classes taken in each year were essentially independent. When heterozygosities were compared within each year for these age and sex classes the significance of the difference in heterozygosity was seen to increase, particularly if cases where either sample was small (< 20) were excluded (Table 5.12).

Table 5.12

Comparison between sites of heterozygosities at each locus, within adult male, adult female and juvenile classes in each year.

Comparison	Heterozygosity higher at		G-test
	SB	BR	
All[‡]	18	34	G = 5.00, $p < 0.05$
For N ⩾ 20 only	9	28	G = 10.24, $p < 0.005$

[‡] excludes 2 instances where equal at each site

The preliminary statistics for differences between the nestling

samples (Table 5.10) suggest electromorph frequency differences for IDHC, PEPT, GP1 and EST2. If the nestlings were independent, the differences would be significant below the 1% level for IDHC, PEPT and GP1 (IDHC: $\underline{X}^2_1 = 8.38$, $\underline{p} = 0.004$; PEPT: $\underline{X}^2_3 = 12.06$, $\underline{p} = 0.007$; GP1: $\underline{X}^2_2 = 53.73$, $\underline{p} < 0.0001$) and close to this level for EST2 ($\underline{d} = 2.56$, $0.02 > \underline{p} > 0.01$). The GP1 differences alone would appear to be accompanied by a large difference in heterozygosity, though heterozygosity estimates are again higher at Brackenhurst for 5 of the 6 comparisons (Table 5.11). In view of the effects of non-independence among nestlings (see above), only those loci suggesting large allele frequency differences have been investigated further.

A better estimate for heterogeneity between sites for nestling allele frequencies has been obtained by summing the within-year estimates weighted for relationships as presented above for IDHC and GP1 (Tables 5.7 and 5.9) and here for PEPT (Table 5.13). Summing the weighted within-year totals does not allow for non-independence among years, which will occur due to some parents contributing to the offspring samples in more than one year. Since many parents were not identified, it would in any case be impossible to allow totally for non-independence. To treat all nestlings sampled at the same nest among years as full sibs, as was done for weightings within years, would be unduly conservative and wasteful of information.

The summed scores are given in Table 5.14, where the totals for each site are compared. No significant differences were found for IDHC and PEPT, but the result for GP1 remained highly significant. The degree of heterogeneity for GP1 was higher than might reasonably be accounted for by any lack of independence. The difference was due to a much higher frequency of $\underline{GP1^A}$ and $\underline{GP1^C}$ at Brackenhurst.

Table 5.13

Weighted PEPT allele scores for nestlings within years at each site. Refer to Table 5.7 and text for explanation.

Site	Year	Alleles				
		A	B	C	D	E
BR	1980	0.00	0.00	0.00	17.40	0.00
	1981	0.00	3.70	2.57	92.44	0.75
	1982	0.00	4.00	4.91	122.38	2.81
SB	1980	0.00	1.11	0.89	36.38	0.00
	1981	0.00	4.00	1.40	133.83	0.00
	1982	1.18	1.92	3.46	163.64	0.50

Table 5.14

Comparison between sites of nestling allele frequencies for IDHC, PEPT and GP1. The allele scores were obtained by weighting within years (Tables 5.7, 5.9 and 5.13) and summing across years.

Locus	Site	Alleles					Comparison
		A	B	C	D	E	
PEPT	BR	0.00	7.70	7.48	232.22	3.56	$X^2_4 = 6.34$
	SB	1.18	7.03	5.75	333.85	0.50	$p = 0.175$
IDHC	BR	172.46	78.38				$X^2_1 = 3.18$
	SB	263.12	86.13				$p = 0.075$
GP1	BR	65.00	176.40	9.65			$X^2_2 = 20.28$
	SB	42.53	297.64	8.13			$p < 0.0001$

The frequencies of these alleles were also higher at Brackenhurst for adults and juveniles, though less significantly ($p < 0.07$: see above).

In conclusion, then, there were pronounced differences in allele and genotype frequencies at the two study sites. These differences were statistically significant when all testable loci were considered together, and individually for PEPT. Heterozygosity per locus was, overall, significantly higher for the adult/juvenile class at Brackenhurst, and again for PEPT alone in locus by locus tests. Comparisons between nestlings were less powerful, and a very significant difference in allele frequency was found only for GPI. Whether or not statistically significant, the differences between sites for allele frequencies and heterozygosity per locus were generally consistent for the adult/juvenile and nestling age classes.

5.3.7 Linkage Disequilibria

It was important to check for the nonrandom association of alleles between loci. Only two locus digenic components, generally referred to as linkage or gametic phase disequilibria, were investigated. It has been shown above that samples among years and age classes within sites were generally homogeneous. Pooled samples were therefore analysed initially, though the parent-offspring and sib correlations might contribute to spurious significance for the statistical tests of disequilibria. This lack of independence is not likely to counteract results concerning genuine associations, and so providing that independence is ultimately considered the initial computations are minimised.

Electrophoretic data do not allow the complete identification

of genotypes as the two kinds of heterozygote (AB/ab and Ab/aB) cannot be distinguished. Hill (1974) has provided an iterative maximum-likelihood method for the estimation of disequilibrium in such cases. Cockerham and Weir (1977) pointed out that the usual measure of disequilibrium, $\underline{D}$ consists of within and between individual components ($\underline{D}_w$ and $\underline{D}_b$ respectively) such that:

$$D = D_w + 2D_b$$

Hill's method assumes $\underline{D}_b = 0$ (Weir & Cockerham 1979). Provided that the sampled populations are randomly mating, the between individual component should equal zero. There was little evidence for departure from Hardy Weinberg expectations (see above), so that Hill's method could be applied (Weir & Cockerham 1979).

The commonest allele at each locus was at a high frequency (always above 70% for loci with 3 or more alleles). As with similar studies (e.g. Langley, Tobari and Kojima 1974), alleles other than the commonest were pooled to form a single class, and thus only one 3x3 table of genotypes was examined for each test. It was not possible to distinguish alleles in coupling from those in repulsion. Conventionally, the commonest allele at each locus is regarded as being in coupling and the alternative class as in repulsion (Langley and Crow 1974). Positive values of $\hat{\underline{D}}$ refer to an excess over binomial expectation of coupling gametes and negative values refer to a deficiency.

The values for $\hat{\underline{D}}$ obtained are shown in Table 5.15. (EST2 was excluded because of the problems with null alleles - see Section 5.3.2 - although there was no evidence of disequilibria when EST2 phenotypes were analysed as though they were genotypes.) The estimates of $\hat{\underline{D}}$ were tested for significance using both the log likelihood test of Hill (1974) and the chi-square test of Weir

Table 5.15

Coefficient of linkage disequilibrium ($\hat{D}$) for pairs of loci in each study population. Refer to text (Section 5.3.7) for details.

Loci	Site	
	SB	BR
6PGD - PEPD3	-0.0012	-0.0012
6PGD - PEPD2	+0.0004	+0.0030*
6PGD - IDHC	-0.0025	+0.0057
6PGD - PEPT	-0.0012	+0.0044*
6PGD - GP1	-0.0011	+0.0020
PEPD3- PEPD2	+0.0002 †	- †
PEPD3 - IDHC	+0.0005	+0.0012
PEPD3 - PEPT	+0.0009	+0.0089**
PEPD3 - GP1	+0.0058	+0.0049
PEPD2 - IDHC	+0.0052	+0.0027
PEPD2 - PEPT	+0.0009	-0.0010
PEPD2 - GP1	-0.0004	-0.0089*
IDHC - PEPT	-0.0032	-0.0005
IDHC - GP1	-0.0082	-0.0067
PEPT - GP1	-0.0000	+0.0033

* $p < 0.05$ ** $p < 0.01$

† all cases involving PEPD2^{A} alleles omitted (see Section 3.1.3), leaving only 7 informative heterozygotes in SB sample and none in BR sample.

(1979). Weir's method has the potential advantage of not requiring the assumption of Hardy Weinberg equilibrium at the tested loci. In view of the absence of significant departure from Hardy Weinberg ratios it was not surprising that both tests were found to be consistent. Four pairs of loci in the Brackenhurst sample were found to produce statistically significant estimates for $\hat{D}$. These four pairs were reexamined in the combined Brackenhurst adult/juvenile class only. Only one remained significant, and one other approached significance (PEPD2-GP1: $\hat{D} = -0.0101$, $X^2_1 = 3.86$; PEPD3-PEPT: $\hat{D} = +0.0085$, $X^2_1 = 3.73$). No attempt is being made here to correct the nestling samples for parent-offspring and sib-sib correlations. There are, however, known to be several sampled nestlings from different clutches where the parents were not sampled which, if included, would increase the absolute value of estimates of D.

5.4 Discussion

Disequilibria are potentially valuable indicators of nonrandom genetic processes. Two kinds of disequilibria have been investigated here: those among alleles at single loci (Hardy Weinberg disequilibria) and those for pairs of alleles between pairs of loci (linkage disequilibria).

5.4.1 Single locus disequilibria

Hardy Weinberg disequilibria may result from a variety of demographic and selective processes. Some of these are expected to result specifically in increased homozygosity relative to Hardy Weinberg expectations and are tested for using the inbreeding model

(Wright 1922, 1951). Statistical tests for the significance of F are of very low power, so much so that Ward and Sing (1970) were led from theoretical considerations to conclude that the level of inbreeding required to produce a significant deviation from random expectation is so great that it "would be illogical to attribute" significant deviations to inbreeding. Haber (1980) criticised Ward and Sing's reasoning, but came to the same conclusion. Consequently, the lack of any significant deviation at any individual locus is neither surprising nor particularly informative. However, the estimates for the inbreeding coefficient, $\underline{F}$, obtained from the genotypic distributions at different loci are essentially independent, and some of the shortcomings of tests made at individual loci may be potentially overcome if there is agreement among the different $\hat{\underline{F}}$ values. Inspection of the values obtained (Table 5.5) clearly indicates a lack of any concordance among the independent estimates, which appear to fluctuate about a mean not significantly different from zero in both study populations.

5.4.2 Linkage Disequilibria

Two pairs of loci (PEPT-PEPD3 and GP1-PEPD2) showed significant linkage disequilibrium at one site (see above). Breeding data relevant to these pairs of loci was obtained for PEPT/PEPD3 only. One family of 3 offspring showed at least one 'recombination' event and the two loci are not therefore likely to be closely linked. Such genetic disequilibrium between apparently unlinked loci is perhaps surprising. In view of the number of tests conducted (29), two results each with the observed significance must be of dubious value. The power of the tests is, however, a function of the sample

size and allele frequencies at each locus and in 19 of the 29 tests will be lower than in the one producing the most significant value for $\hat{D}$. Following Weir (1979), if p and q are the frequencies of alleles at the two loci such that

$$(1-p) \leqslant (1-q) \leqslant q \leqslant p$$

then

$$D_{min} = -(1-p)(1-q) \leqslant D \leqslant p(1-q) = D_{max}$$

Thus for PEPD2-GP1 in the adult/juvenile class at Brackenhurst, $\hat{D}$ was found to equal D_{min}. This is consistent with a total absence of coupling gametes of the rarer class in Brackenhurst adults and juveniles. For each two locus gametotype there are three genotypic classes which allow their unambiguous detection. For the rarer coupling class, the number expected to be observed is given by

$$4n(1-p)(1-q)\{p(1-q) + q(1-p) + (1-p)(1-q)\}$$

where n is the number of sampled individuals. For PEPD2-GP1, this random expectation was 4.7 for Brackenhurst adults (none was observed), and 7.1 for Brackenhurst pulli (2 were observed). In contrast, 8 were observed at Sutton Bonington (including 6 adults) against an expectation of 5.6. Thus if the observed disequilibrium at Brackenhurst is a true reflection of nonrandom processes in the population, then those processes will be operating differently at the two sites. Such differences are not unexpected by population genetic theory (e.g. Hedrick *et al.* 1976), particularly as there is a large difference in allele frequency at GP1 (see above). A comparison of $\hat{D}$ (= 0.0085) with D_{max} (= 0.0909) for PEPD3-PEPT reveals that the disequilibrium for this pair of loci, in this case due to an excess of coupling gametes, is far less extreme.

The nonrandom processes producing the estimates of disequilibrium could include selection or population subdivision. Genetic or other effects leading to modification of the gene products at two or more loci may lead to apparent disequilibria (as observed in studies of Drosophila, G.B. Johnson, pers. comm.). Such modification would, however, have been detected in the analysis of broods (see Section 3.2.1) and none was. Further, this effect might have been expected to apply to samples from both sites.

Lewontin (1974) has shown from a theoretical standpoint that fitness differences among genotypes may lead to disequilibria even between unlinked loci. For one locus (PEPD3 or GP1) in each pair showing disequilibria gametic selection has been suggested as a possible explanation for significant deficiencies in the transmission of the rarer alleles (Chapter 3). There was, however, no evidence for gametic selection at either PEPT or PEPD2, and no informative matings were available to test for any effect of PEPD2 genotype upon GP1 transmission, and insufficient to test for any effect of PEPT upon PEPD3. Thus there was no corroborating evidence for selection as a cause of disequilibria here.

Although there was no significant deviation from Hardy Weinberg ratios, there was some evidence for the nonrandom distribution of genotypes amongst loci in those families examined for genetic incompatibilities (Section 4.3.3). This might point to the existence of multiple locus disequilibria, which have not been tested for, possibly resulting from a degree of subdivision or non-random mating within the samples. The extent of the differences in genic distribution between the two study populations demonstrates that significant differentiation is possible, but a comparison of samples taken within the two separate parts of the Brackenhurst site ('G'

and 'H' in Figure 2.2) did not provide any evidence of detectable spatial differences in gene frequency at individual loci.

From an expectation that epistatic interactions among loci will be extensive, a theoretical prediction has been made that disequilibria among both linked and unlinked loci should be widespread (Franklin and Lewontin 1970, Lewontin 1974). Evidence has been obtained against most of the loci studied here being linked (Chapter 3). The number of loci was small, but the results were not at variance with studies of natural populations of Drosophila, where no convincing evidence of disequilibria among unlinked loci has been found (Loukas, Krimbas and Morgan 1980). Indeed, such empirical studies of Drosophila suggest disequilibria among allozyme loci occur at detectable strengths only between very closely linked loci or loci associated with inversions (reviewed in Loukas, Krimbas and Vergini 1979).

The PEPD2-GP1 locus pair was one of the few for which no relevant breeding data were available; in view of the disequilibrium result obtained, whether these loci are linked will ultimately be of particular interest. In studies of Drosophila a conclusion that epistatic fitness effects are the cause of those disequilibria observed among loci (excluding those in inversions) relies largely on the finding of the same disequilibria in several populations (Loukas, Krimbas and Morgan 1980). As there was no evidence for disequilibria at Sutton Bonington, the disequilibrium for PEPT-PEPD3 at Brackenhurst is thought to be most probably due to sampling or demographic effects. Disequilibria might be expected to be generated more frequently by these processes in vertebrate populations. Natural vertebrate populations will be far more structured than those of Drosophila. No data concerning both PEPD2

and GP1 or both PEPD3 and PEPT is available for large samples from populations other than those studied here.

5.4.3 Comparisons of Genic Distributions

Of the comparisons made among various classes, substantial differences in genic distributions were found only for comparisons between the two study populations. It must be pointed out that these particular tests would have had the highest statistical power of those carried out at any particular locus as the maximum number of sampled individuals was included in these tests. Nevertheless, no other trends were suggested except for the tendency of Sutton Bonington adults to have a lower frequency of the commonest allele at each locus than did nestlings (7/7 comparisons, $p = 0.0156$, 2-tailed binomial). This was not the case at Brackenhurst, however, and the expectation that the frequency difference at Sutton Bonington would reflect increasing heterozygosity in adults was not supported by the data.

There were substantial differences in heterozygosity between the two sites. From stochastic theory, heterozygosity is expected to be maintained at a higher level in larger populations (Kimura and Crow 1964; cf. Section 1.2). Thus it might be predicted that the size of the sampled population at Brackenhurst will exceed that at Sutton Bonington. Many of the birds sampled at each site were retrapped on at least one occasion. Though far fewer adults were marked at Brackenhurst than at Sutton Bonington (BR: 239; SB: 453), the overall retrap rate at Brackenhurst was much higher than at Sutton Bonington (Table 5.16). It is difficult to assess the relative randomness of retrapping at each site, but it is extremely unlikely that methodological differences could account for the

observed difference in recapture rate, and in view of the smaller numbers of birds marked at Brackenhurst the sampled population at that site would appear to be much smaller than at Sutton Bonington. Thus the differences in heterozygosity would not appear to be attributable to sampled population size differences. The size of the sampled study populations may not, however, be a useful guide to

Table 5.16

Details of ringing totals and retraps for adults marked at each study site.

Site	Year	Number marked	% Retrapped per Site per Year		
			1980	1981	1982
SB	1980	116	15.5	28.4	1.7
	1981	238		8.0	6.7
	1982	99			10.1
BR	1980	129	13.2	10.1	3.1
	1981	46		17.4	39.1
	1982	64			25.0

the effective population size. It is conceivable that the arbitrarily defined sampled populations are actually part of much larger, reasonably panmictic groupings and that the pattern of movements differs at the two sites. For example, if the area used

for sampling at Sutton Bonington was more attractive to birds from the population as a whole for feeding than the equivalent area at Brackenhurst, then the rate of recapture at Sutton Bonington might be lower than at Brackenhurst even if the effective population size at Brackenhurst was higher. It is quite possible that the main trapping area at Sutton Bonington, a mill building containing permanently accessible grain, attracted sparrows from greater distances than did the trapping areas at Brackenhurst.

In view of the observation of distorted segregation ratios for GP1 and PEPD3, the comparison of allele frequencies among age classes for these loci was of particular potential interest. As shown above (Section 5.3.5), there was no evidence of a change in gene frequency among age classes for these loci. Thus no mechanism to balance the effects of the distorted segregations has been discovered (cf. Chapter 3).

5.5 Conclusions

The separate populations studied were not found to disagree with the null hypothesis that they each represented an essentially random breeding unit. In view of the power of tests for disequilibria at single loci, any other result would have been surprising. A consideration of several essentially independent estimates of inbreeding as found for different loci was potentially more powerful, but no suggestion of significant inbreeding (or outbreeding) resulted although the average value in each study population was positive. There was some evidence for disequilibria between pairs of loci at Brackenhurst, particularly for PEPD3-PEPT, but this was thought most likely to be due to sampling error. It

was unfortunate that informative breeding data regarding linkage was not available for all pairs of loci, particularly in the case of PEPD2-GP1, where possible disequilibrium was found. Considerations of digenic disequilibria in both populations and a comparison of genotypic distributions between the separate parts of one study site did not provide any evidence for structuring within populations. No conclusions concerning the possible existence of multigenic disequilibria can be made here, but in view of the suggestion of multigenic effects found in a study of exclusion probabilities (Chapter 4), the analysis should in the future be extended to a consideration of such disequilibria.

In the comparisons of genic distributions involving nestlings the confounding effects of sib correlations and the importance of making allowance for them was emphasised. Comparisons of the distributions of genotypes and alleles among sex, age and year classes did not provide any evidence for selection; the factors presumed to balance the effects of distorted segregation ratios found for GP1 and PEPD3 remain unknown. There were large, unexpected, differences in allele and genotype frequencies between the two study populations, with one population having a significantly higher level of heterozygosity than the other. The available ecological and mark-recapture data were inadequate to allow a comparison of effective population sizes at the study sites, but in view of the observed differences in heterozygosity a more intensive study of this aspect could be worthwhile.

5.6 Summary

Genotype and allele frequencies were presented for 7 polymorphic protein loci in each age, sex and sampling year class at each of two study sites. The samples were not found to depart from Hardy Weinberg equilibrium, and there was no evidence for significant inbreeding within sites. There were no differences in allelic distributions between the sexes or among years for adults within the populations. No differences were found among age groups or nestling year classes when allowance was made for sib correlations. Heterozygosities were higher at Brackenhurst than at Sutton Bonington for most loci, and the overall difference was significant. There was a particularly large difference in allele frequencies between nestlings in the separate populations for GPl.

There was evidence for digenic gametic disequilibrium at Brackenhurst for some combinations of loci, but this was thought to be most likely due to sampling effects. No evidence was obtained for population substructuring within the populations, or for selection at any locus.

CHAPTER 6

ASPECTS OF MATING

6.1. Introduction

Most of theoretical population genetics assumes that populations are randomly mating. It is, however, predicted that the ability to choose a mate of high quality will be adaptive. If this chosen quality is heritable, then mate choice will lead to nonrandom mating at those loci, and any marker loci linked to them, contributing to those phenotypic characteristics used in the assessment of quality. If mates are chosen with respect to their level of relatedness, then nonrandom mating will occur at all genomic loci. If mates are chosen with respect to physiological, morphological or behavioural characteristics, then only those loci (or loci tightly bound to them) contributing directly to the phenotype are expected to be affected.

Among diploid species, such nonrandom mating may result in the detection of deviations of genotypic proportions from those that would be expected if gametes were associating randomly. However, such deviations from Hardy Weinberg equilibrium are not necessarily attributable to active mate choice as they may also result from gene-flow, selection or undetected population substructure. Electrophoretic studies routinely include the testing of genotypic ratios for agreement with Hardy Weinberg expectations. In general, repeatably demonstrable deviations have been observed only rarely. In particular, no such deviations have been observed in house sparrow populations.

The Hardy Weinberg test is, in any case, not a powerful one

(see Section 5.4.1). The analysis of mate pairs' genotypes is potentially much more informative as to whether mating is random. Firstly, mate pair data contains much more information, in the statistical sense, than population data. Tests for inbreeding are three times more powerful when mate pair data are used than for population data (Yasuda 1969, Cockerham 1973). Such data are, however, generally more difficult to collect than data from random individuals. Secondly, there are fewer confounding effects such as natural and sexual selection and migration which might apply differentially between the successful breeders and the remainder of the population. Analyses based solely on population data could be highly misleading. For example, it is theoretically possible for mating to be absolutely assortative without any deviation from Hardy Weinberg ratios resulting.

In this chapter I examine the electrophoretic data for mate pairs in a number of ways. Firstly, mate pair frequencies are compared with random expectation. Secondly, using analysis of variance arguments (Cockerham 1973) the correlations of genes within and between mates are estimated. These estimates are analogous to inbreeding coefficients (Wright 1965). Thirdly, the results of an analysis using coefficients of genetic similarity (Rogers 1972) are compared with those from the genetic correlation analysis. Metric data for mate pairs are also examined to test for the possibility of nonrandom mating.

6.2. Methods

The field and laboratory methods (Chapter 2) and protein polymorphisms (Chapter 3) have been described previously. Mate

pairs were selected only from complete families where the offspring were also sampled. Where both adults remained together for more than one clutch only the first mating was included. Matings where one parent was replaced were included. Where measurements were made on separate occasions these were averaged. The two study populations were analysed separately. Sample sizes are shown below together with the number of individuals involved.

Table 6.1

Number of mate pairs included in mating analyses, with the number of different individuals comprising the mate pairs.

	Site	
	BR	SB
Different Matings	33	66
Different males	29	51
Different females	31	55

All 6 codominant loci were included in the analysis for each mate pair except where PEPD2 was heterozygous in which case PEPD3 was excluded (Chapter 3).

6.3 Results

6.3.1 Mating Frequencies

The observed mating type frequencies are shown in Table 6.2. To simplify the analysis, the sexes have been ignored and all alleles other than the commonest at a locus have been pooled (as R). This provides six mating types per locus.

Also shown in Table 6.2 are the binomial expectations (see Table 6.3) for each mating class in each population for each of the three models:

E1 - random mating among family subsample.

E2 - random mating among family subsample, allele frequencies equal in each sex.

E3 - random mating among all sampled adults, allele frequencies equal in each sex.

The expectations based on the different models were not found to differ significantly. The assumption of equal allele frequencies in the sexes in model (ii) can be seen to have had minimal effect upon expectations; only IDHC at Sutton Bonington showed absolute differences exceeding 0.1. The same assumption in model (iii) was found to make even less difference when compared with expectation based upon separate allele frequencies for each sex (not shown: all changes < 0.1).

The observed mating type frequencies did not, at any locus or population, differ significantly from expectation.

6.3.2. Genic Correlations

Cockerham (1973) described the components of total genic variance (σ^2) for genes within individuals (σ^2_w), genes between

Table 6.2

Observed and expected mating type frequences for each locus at each site. The models on which expectations were based are:

E1 - random mating among males and females comprising mate pair samples.

E2 - as for E1, but assuming allele frequencies equal in each sex.

E3 - random mating among males and females in entire adult samples.

R = all alleles other than the commonest combined.

Table 6.2

Locus	Mating	Sample							
		BR				SB			
		Obs	E1	E2	E3	Obs	E1	E2	E3
6PGD	BB x BB	29	29.2	29.2	27.3	63	63.0	63.1	59.8
	BB x BR	4	3.6	3.6	5.3	3	2.9	2.9	6.0
	BB x RR	0	0.0	0.0	0.1	0	0.0	0.0	0.1
	BR x BR	0	0.1	0.1	0.2	0	0.0	0.0	0.1
	BR x RR	0	0.0	0.0	0.0	0	0.0	0.0	0.0
	RR x RR	0	0.0	0.0	0.0	0	0.0	0.0	0.0
PEPD3	BB x BB	17	18.5	18.5	21.0	42	42.1	42.1	40.6
	BB x BR	11	8.1	8.1	6.2	11	10.8	10.8	12.0
	BB x RR	0	0.4	0.4	0.2	0	0.3	0.3	0.4
	BR x BR	0	0.9	0.9	0.5	1	0.7	0.7	0.9
	BR x RR	0	0.1	0.1	0.0	0	0.0	0.0	0.1
	RR x RR	0	0.0	0.0	0.0	0	0 0	0.0	0 0
PEPD2	BB x BB	28	27.4	27.4	27.3	51	52.2	52.2	50 0
	BB x BR	4	5.2	5.2	5.3	15	12.6	12.6	14.4
	BB x RR	0	0.1	0.1	0.1	0	0.4	0.4	0.5
	BR x BR	1	0.2	0.2	0.2	0	0.7	0.8	1.0
	BR x RR	0	0.0	0.0	0.0	0	0.0	0.0	0.1
	RR x RR	0	0.0	0.0	0.0	0	0.0,	0.0	0.0
IDHC	AA x AA	8	7.4	7.5	7.7	19	18.4	18.9	18.6
	AA x AB	13	13.5	13.4	13.5	28	28.3	27.7	27.8
	AA x BB	2	3.1	3.0	3.0	10	6.3	5.1	5.3
	AB x AB	6	6.0	6.1	5.9	4	9.2	10.2	10.2
	AB x BB	4	2.7	2.7	2.6	5	3.5	3.8	3.8
	BB x BB	0	0.3	0.3	0.3	0	0.3	0.3	0.3

Table 6.2 contd.

Locus	Mating	Sample							
		BR				SB			
		Obs	E1	E2	E3	Obs	E1	E2	E3
PEPT	DD x DD	23	24.1	24.1	22.0	48	48.9	48.9	51.1
	DD x DR	10	7.9	7.9	9.4	17	15.3	15.3	13.5
	DD x RR	0	0.3	0.3	0.5	0	0.6	0.6	0.5
	DR x DR	0	0.6	0.6	1.0	1	1.1	1.1	0.9
	DR x RR	0	0.1	0.1	0.1	0	1.1	1.1	0.1
	RR x RR	0	0.0	0.0	0.0	0	0.0	0.0	0.0
GP1	BB x BB	10	12.7	12.7	12.4	31	28.4	28.4	28.2
	BB x BR	18	13.7	13.7	13.8	22	26.8	26.8	26.7
	BB x RR	1	1.9	1.9	1.9	5	3.3	3.3	3.2
	BR x BR	4	3.7	3.7	3.8	7	6.1	6.1	6.3
	BR x RR	0	1.0	1.0	1.0	0	1.4	1.4	1.5
	RR x RR	0	0.1	0.1	0.1	1	0.1	0.1	0.1

Table 6.3

Expressions for expected proportions of mating types.

Mating Type *	Expected Proportion **
AA x AA	$p_m^2 \, p_f^2$
AA x AR	$2p_m(1-p_m)p_f^2+2p_f(1-p_f)p_m^2$
AA x RR	$(1-p_m)^2p_f^2+(1-p_f)^2p_m^2$
AR x AR	$4p_mp_f(1-p_m)(1-p_f)$
AR x RR	$2p_f(1-p_f)(1-p_m)^2+2p_m(1-p_m)(1-p_f)^2$
RR x RR	$(1-p_m)^2(1-p_f)^2$

* A = commonest allele, R = rest combined

** p_m = frequency of A in males

p_f = frequency of A in females

mates ($\underline{\sigma}^2_v$) and between mate pairs ($\underline{\sigma}^2_m$). From these components the correlations of genes within individuals ($\underline{F}$) and of genes between mates ($\underline{\theta}$) are estimated as:

$$F = (\sigma^2_m + \sigma^2_v)/\sigma^2$$

and $$\theta = \sigma^2_m/\sigma^2$$

Cockerham (1973) provides χ^2 tests for the following hypotheses:

$$H_1 : F = 0$$
$$H_2 : \theta = 0$$
$$H_3 : \theta = F = 0$$
$$H_4 : F = \theta$$

(Note that Cockerham's (1973) equation (14) is misprinted - refer to equation (11).)

If $\underline{H}_4$ is accepted then a pooled estimator, $\hat{\underline{\alpha}}$, may be obtained as

$$\hat{\alpha} = (\hat{F} + 2\,\hat{\theta})/3$$

Thus $\underline{H}_3$ is the test of $\underline{\alpha} = 0$, and is conditional upon acceptance of $\underline{H}_4$. Cockerham's $\underline{F}$ is equivalent to the fixation index of Wright (1969). If $\underline{F} = \underline{\theta}$, the population is at equilibrium, and if inbreeding is the only factor operating, then $\underline{\theta}$ will be the best estimate of the inbreeding coefficient.

The estimates $\hat{F}$, $\hat{\theta}$ and $\hat{\alpha}$ as obtained following the procedure of Cockerham (1973), are given in Table 6.4. All four hypotheses outlined above were tested for each set of estimates. Population averages for the estimates $\hat{F}$, $\hat{\theta}$ and $\hat{\alpha}$ were obtained by summing the variance component estimates obtained separately for each of the six loci. This is the procedure as recommended for $\bar{\theta}$ by B.S. Weir (pers. comm.) and Weir and Cockerham (1984).

Thus for n loci:

$$\bar{\theta} = \sum_{i=1}^{n} \sigma^2_{m_i} / \sum_{i=1}^{n} \sigma^2_i$$

$$\bar{F} = \sum_{i=1}^{n} (\sigma^2_{m_i} + \sigma^2_{v_i}) / \sum_{i=1}^{n} \sigma^2_i$$

$$\bar{\alpha} = \sum_{i=1}^{n} (3\sigma^2_{m_i} + \sigma^2_{v_i}) / \sum_{i=1}^{n} \sigma^2_i$$

Numerical estimates of the variances of these combined loci estimators were found by the jackknife procedure, jackknifing over loci by omitting each in turn (Reynolds, Weir and Cockerham 1983; Weir and Cockerham 1984). None of the combined estimators was found to differ significantly from 0 (Table 6.4).

Amongst the tests at individual loci, significant departures were observed for IDHC matings at Sutton Bonington for $H_1 : F = 0$ ($\chi^2_1 = 6.62$, $p < 0.01$) and $H_4 : F = \theta$ ($\chi^2_1 = 5.33$, $p < 0.01$). The power of the tests in this particular instance will be one of the highest for the data set presented, as they applied to the larger sample and one of the more heterozygous loci. In an assessment of whether they are spurious or meaningful, they should not therefore

be regarded as simply representing 2 tests out of a total 36. As, however, a significant positive value was not obtained for $\underline{\theta}$, the observed departures cannot be attributed to inbreeding. Further, the data for the other loci do not support any hypothesis regarding inbreeding or any other form of nonrandom mating. In particular, there was no concordance between the results for IDHC in the two separate populations. A closer examination of the mate pair data for IDHC at Sutton Bonington revealed a significant difference in the genotypic distributions for the two sexes:

Table 6.5

IDHC genotypes of mate pairs at SB

Sex	Genotype			
	AA	AB	BB	
Males	44	19	3	
Females	32	22	12	
				$G_2 = 7.90$
				$0.025>p>0.010$

Cockerham's formulae assume that the allele frequencies in the two sexes are the same; the anomalous result for IDHC at Sutton Bonington resulted from this condition not being met. The apparent difference between the sexes in the mate pair subsample was itself an effect of sampling. The 12 $\underline{IDHC^{B/B}}$ female genotypes in different mate pairs are actually contributed by only 7 different females. 2

Table 6.4

Estimates of correlations between genes in mate pairs. Combined estimates are shown ± 1SE.

Locus	Site	$\hat{F}$	$\hat{\theta}$	$\hat{\alpha}$
6PGD	BR	-0.0503	-0.0503	-0.0503
	SB	-0.0066	-0.0062	-0.0063
PEPD3	BR	-0.1021	-0.1021	-0.1021
	SB	-0.0589	0.0229	-0.0044
PEPD2	BR	-0.0382	0.1352	0.0774
	SB	-0.0567	-0.0567	-0.0567
IDHC	BR	-0.0867	0.0526	0.0062
	SB	0.2132*	-0.0939*	(0.0085)*
PEPT	BR	-0.0756	-0.0756	-0.0756
	SB	-0.0738	-0.0172	-0.0361
GP1	BR	-0.2409	-0.1455	-0.1773
	SB	0.1157	0.0421	0.0667
Combined	BR	-0.0982 ±0.0673	-0.0358 ±0.0875	-0.0498 ±0.0735
	SB	0.0897 ±0.0972	-0.0277 ±0.0420	0.0114 ±0.0354

*significant departures from null hypotheses – refer to text.

each mated with 2 different males and 1 other was observed to mate with a total of 4 different males during the course of the study (a unique example).

6.3.3 Similarity Coefficients

The use of coefficients of genetic identity allows an alternative approach to an assessment of the genetic similarity of mates. The most frequently used indices for population comparisons are Rogers' S (Rogers 1972) and Nei's I (Nei 1972). Schwartz and Armitage (1983) applied both coefficients to electrophoretic data for individuals from colonies of the yellow-bellied marmot, Marmota flaviventris. They assessed the usefulness of the coefficients for inferring relatedness by comparing the values obtained for each pair (not just mates) of individuals in each colony whose pedigrees were known, with the coefficient of relatedness for each pair as determined by path analysis (Falconer 1981).

Rogers' S for a pair of individuals is found as:

$$S = 1-\{(\sum_{i=1}^{m}[P_{i.x} - P_{i.y}]^2)/2\}^{0.5}$$

where $P_{i.x}$ is the frequency of the ith allele in individual x and m is the number of alleles. As, for an individual, $m = 2$ it follows that $S = 1.0$, 0.5 or 0.0 depending on whether the pair have 2, 1 or 0 alleles respectively in common. Nei's index for a pair of individuals is of the form:

$$I = \sum_{i=1}^{m} p_{i.x}\ p_{i.y} \Big/ \left(\sum_{i=1}^{m} p_{i.x}^2 \sum_{i=1}^{m} p_{i.y}^2\right)^{0.5}$$

For a pair's genotypes the values of I are the same as for S except in the case of the pairing of a heterozygote with a homozygote having one allele in common. Nei's index effectively treats such a pairing ($I = \sqrt{0.5}$) as being more similar than one consisting of two heterozygotes having one allele in common ($I = 0.5$). All possible types of pairs, using mating terminology, are shown with appropriate S and I values in Table 6.5. Schwartz and Armitage (1983) found that I was less correlated with relatedness than S. They concluded that whilst S could not usefully be applied quantitatively to predict relatedness, it had on average a significantly larger value for related as compared with unrelated pairs. Whilst Schwartz and Armitage were particularly concerned with measures of intracolony relatedness, their conclusions should apply equally to mate pairs.

As Rogers' S is easily estimated and, unlike the more rigorous measures of genic correlations of mate pairs (see above), as it can be found for mate pairs considered individually, it is of interest to know if estimates of S are empirically related to estimates of those correlations. The values that S may take are ultimately constrained by the allele frequencies at a locus. As the allele frequencies vary among loci, raw individual estimates of S from different loci cannot be analysed together. Differences among loci would reflect differences in allele frequency more than any differences in estimates for genic correlations. A relative measure of deviation, d, of mean observed values for mate pairs, $\hat{S}$, from that expected

under a hypothesis of random mating, $\underline{S}_e$, is therefore proposed:

$$d = (\hat{S} - S_e)/(1 - S_e)$$

$\underline{d} = 1$ when all mates are identical and $\underline{d} = 0$ when mating conforms with the assumptions made in the estimation of $\underline{S}_e$.

Values of $\underline{d}$ were found for each locus and study population (Table 6.6). $\underline{S}_e$ was calculated from the observed allele frequencies in the mate pair samples, treating the mate pairs as independent and assuming that the allele frequencies were equal in each sex. The same conditions applied in the estimation of genic correlations (Section 6.3.2). No allelic classes were combined when calculating $\hat{\underline{S}}$ or $\underline{S}_e$. Correlation coefficients were calculated between these 12 values for $\underline{d}$ and the appropriate estimates of the genic correlation coefficients, $\hat{\underline{\theta}}$, $\hat{\underline{F}}$, and $\hat{\underline{\alpha}}$ (from Table 6.4). A significant correlation was obtained between $\hat{\underline{d}}$ and $\hat{\underline{\theta}}$ only ($\underline{r} = 0.661$, $\underline{p} = 0.019$).

Thus in the samples studied here, there appears to be a strong relationship between a measure of the deviation of estimates of the coefficient of genetic similarity $\underline{S}$ from random expectation and estimates of the correlation $\underline{\theta}$ of genes between mates.

6.3.4 Assortative Mating

The analyses of protein polymorphisms just described provided no evidence for assortative mating. Assortative mating with respect to such polymorphisms is in any case particularly unexpected as none of the loci is known to be associated with appropriate potential morphological or physiological mating cues. Morphological characters are in general potential mating cues. One aspect of

Table 6.6

Mean values for Rogers' coefficient of genetic similarity, $\hat{S}$, for mate pairs as determined at each locus in each study population. S_e = value expected from random mating among alleles in mate pair samples. $\hat{d}$ = relative deviation of $\hat{S}$ from S_e (refer to text).

Site	Locus	$\hat{S}$	S_e	$\hat{d}$
BR	6PGD	0.939	0.943	-0.070
	PEPD3	0.833	0.831	0.010
	PEPD2	0.955	0.917	0.457
	IDHC	0.652	0.663	0.034
	PEPT	0.833	0.879	-0.381
	GPl	0.712	0.708	-0.015
SB	6PGD	0.977	0.978	-0.045
	PEPD3	0.886	0.892	-0.055
	PEPD2	0.879	0.899	-0.197
	IDHC	0.583	0.684	-0.320
	PEPT	0.871	0.870	-0.009
	GPl	0.705	0.721	-0.057

these metrical variables was therefore investigated.

Before proceeding, the relationships among different metrics and their distribution between populations were investigated. The house sparrow is dimorphic with respect to size (see Table 6.7) and the sexes were therefore treated separately. One-way analyses of variance revealed small but significant size differences between the male samples (Table 6.7). Males at Brackenhurst were heavier than at Sutton Bonington ($p = 0.024$) and had shorter wings ($p = 0.012$). The weight difference may have been due in part to the trapping time of males at Brackenhurst being on average 1.17h later than at Sutton Bonington ($\bar{x}_{BR} = 11.24h$, $\bar{x}_{SB} = 10.07h$, $F_{1,305} = 7.88$, $p = 0.01$). It was concluded that there may be size differences between the populations, and the two samples were therefore not pooled.

Correlations among the different metrics were investigated (Table 6.8). All 3 tarsus measurements (Chapter 2) were included in the analysis. It was found that where sample sizes were reasonable (>50), TAR2 was the tarsus measurement most correlated with other metrics. The measurement that correlates most is likely to be the most biologically meaningful; hence TAR2 would appear to be the most valuable of the 3 measurements taken in spite of anticipated difficulties with respect to its measurement (see Chapter 2). Unfortunately the measuring of TAR2 commenced at a later stage in the study. Significant values for the product-moment correlation coefficient r, ranging from 0.17-0.55, were observed for almost all combinations of weight and wing, tail and tarsus lengths (taking TAR2 as the best measurement in larger samples), for all four subsamples. The only pair of variables not found to correlate were tail and tarsus, and this again was the case for all four subsamples. Wing and tail lengths were the most correlated in each

Table 6.7

Means ± 1 standard error for size characteristics of adult males and females at each study site, with results (F) for comparisons between sites within sex classes. Sample sizes are shown in parentheses. p = probability ns = not significant.

Sex Class	Variable	Sample		F	p
		SB	BR		
Females	Wing (mm)	75.78±0.14 (169)	75.45±0.16 (133)	2.38	ns
	Weight (g)	26.94±0.17 (168)	27.44±0.19 (134)	3.82	ns
	Tail (mm)	55.76±0.18 (168)	55.35±0.22 (130)	2.06	ns
	TAR1 (mm)	17.62±0.07 (135)	17.44±0.08 (85)	2.79	ns
	TAR2 (mm)	18.41±0.12 (51)	18.41±0.13 (54)	0.00	ns
	TAR3 (mm)	18.59±0.10 (56)	18.44±0.13 (54)	0.85	ns
Males	Wing	78.57±0.12 (201)	77.98±0.18 (102)	7.63	0.012
	Weight	27.19±0.11 (197)	27.68±0.16 (101)	6.41	0.024
	Tail	57.93±0.17 (197)	57.62±0.20 (100)	1.32	ns
	TAR1	17.55±0.06 (157)	17.64±0.09 (63)	0.53	ns
	TAR2	18.52±0.11 (55)	18.61±0.15 (28)	0.25	ns
	TAR3	18.57±0.09 (60)	18.64±0.80 (31)	0.16	ns

Table 6.8

Product-moment correlation coefficients (r) for adult metrics.
N = sample size, p = probability.

Site SB

		Male					Female				
		WEIGHT	TAIL	TARS1	TARS2	TARS3	WEIGHT	TAIL	TARS1	TARS2	TARS3
WING	r	0.257	0.554	0.105	0.346	0.324	0.173	0.339	0.196	0.307	0.213
	N	196	197	155	56	61	167	167	135	51	56
	p	****	****	0.09	**	**	*	****	*	*	0.06
WEIGHT	r		0.252	0.245	0.323	0.422		0.231	0.224	0.315	0.213
	N		192	154	55	60		166	134	49	54
	p		***	**	**	***		**	**	*	0.06
TAIL	r			0.066	0.205	0.192			0.032	0.175	0.128
	N			154	55	60			135	51	56
	p			0.21	0.07	0.07			0.35	0.11	0.17
TARS1	r				0.669	0.661				0.766	0.802
	N				54	54				57	56
	p				****	****				****	****
TARS2	r					0.946					0.897
	N					54					51
	p					****					****

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$
**** $p < 0.0001$

contd.

Table 6.8 contd.

Site BR		Male					Female				
		WEIGHT	TAIL	TARS1	TARS2	TARS3	WEIGHT	TAIL	TARS1	TARS2	TARS3
WING	r	0.213	0.450	0.301	0.089	0.212	0.201	0.443	0.064	0.231	0.196
	N	97	98	62	27	30	132	128	85	54	54
	p	*	****	**	0.33	0.13	*	****	0.28	*	0.08
WEIGHT	r		0.182	0.252	-0.054	0.016		0.183	0.371	0.321	0.335
	N		97	62	28	31		127	85	54	54
	p		*	*	0.39	0.47		*	***	**	**
TAIL	r			0.099	0.306	0.338			0.092	0.101	0.086
	N			62	27	30			83	52	52
	p			0.22	0.06	*			0.20	0.24	0.27
TARS1	r				0.758	0.826				0.797	0.781
	N				28	31				54	54
	p				****	****				****	****
TARS2	r					0.923					0.975
	N					28					54
	p					****					****

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$
**** $p < 0.0001$

of the four groups; tarsus length with weight or wing-length were the second and third most correlated in every group. Thus all four measurements would appear to some degree to reflect the overall size of the individual.

The correlation coefficients for the metrics of mate pairs in each study population were determined (Table 6.9.) As the sample sizes for TAR2 were small, TAR1 has been taken as the tarsus measurement. Significant correlations were found at Sutton Bonington (the larger sample) for weight and tail-length ($p<0.04$), and that for tarsus-length approached significance ($p<0.07$). None of the correlations was significant for the Brackenhurst samples, and none of the correlations for a variable differed significantly between sites.

6.4 Discussion

6.4.1 Genic Correlation and Similarity

The individual locus by locus estimates for the correlation of genes between mates θ, the correlation of genes within individuals (i.e. of uniting gametes) F, and the pooled estimator α, as determined following Cockerham (1973), have been discussed above. Only one previous study of a natural population (of *Peromyscus polionotus*) has to my knowledge used mate pair data to find inbreeding coefficients by Cockerham's method (Foltz 1981). In a large randomly mating population these genic correlations are expected to approximate to zero. The jackknifed within sample estimators were not significantly different from 0 (Table 6.4), nor were $\bar{\theta}$ and $\bar{F}$ significantly different from one another. It is, however, noteworthy that when the twelve independent estimates for F

Table 6.9

Product-moment correlation coefficients (r) for the metrics of mates at each study site.

N = sample size, p = probability

Metric	Mate pair sample					
	SB			BR		
	r	N	p	r	N	p
Wing-length	0.104	66	0.200	0.102	32	0.290
Weight	0.230	63	0.035	-0.080	32	0.331
Tail length	0.232	65	0.031	-0.008	31	0.335
TAR 1	0.261	49	0.068	-0.179	23	0.207

were considered together (2 samples x 6 loci), there were significantly more negative than positive values (10:2, $\underline{p}$ = 0.0386).

There was no significant tendency, however, for $\underline{\theta}$ to be negative. Thus the tendency for $\underline{F}$ to be negative was not attributable to any tendency towards outbreeding. Other causes are therefore more likely; for example, negative values for $\underline{F}$ would be expected if there were an excess of heterozygotes due to selective differences. The values for $\underline{F}$ obtained here were based on a small subsample of adults, and their contributions were partly correlated with the number of successful clutches with which they were associated (as parents which changed mates were re-included in the analysis). The values obtained for $\underline{F}$ from the entire adult samples (Section 5.3.3 and Table 5.5) showed no significant tendency to deviate from zero, and it might then be postulated that successful breeders tend to be more heterozygous than other adults.

The finding that a proposed measure of deviation from genetic similarity, $\underline{d}$, was correlated with $\underline{\theta}$ is potentially useful as the calculation of $\underline{d}$ is relatively convenient, particularly if some combined measure of similarity of mates across several loci is required. The statistical properties of $\underline{d}$ have not, however, been investigated and testing the significance of results must rely on the use of nonparametric tests for trends in values obtained independently.

6.4.2 Assortative Mating

The suggestion of assortative mating at Sutton Bonington for some size variables, particularly tail-length, is of particular interest. There are several examples of assortative or disassortative mating for plumage characteristics in birds. For

example, lesser snow geese, Anser c. caerulescens, mate assortatively with respect to body plumage colour (Cooke et al. 1976) and white-throated sparrows, Zonotrichia albicollis, mate disassortatively with respect to crown stripe colour (Lowther 1961, cf. Thorneycroft 1976). Metric data concerning mate pairs in birds has only occasionally been presented in the literature. Boag and Grant (1978, cf. Boag 1983) obtained data for the medium ground finch, Geospiza fortis, which suggested that assortative mating for size characters, particularly bill size and tarsus length, occurs in some years.

Correlations of size variables have more often been observed in groups other than birds, and have been particularly well documented in humans (Roberts 1977). Most components of size in humans are found to correlate. If size per se was the characteristic with respect to which sparrows were mating assortatively then, on the basis of the strengths of the correlations among different variables (Section 6.3.4 and Table 6.8), wing-length might have been expected to show the largest effect. That tail-length showed the largest effect might point to the possibility that tail-length is itself closely correlated with a particular mating cue. If there is an active assessment of size between potential mates, then tail size might itself have an influence, particularly as males hold their tails in a raised position during solicitation displays. There would not, however, appear to be much advantage in mating attributable to size itself, as there is only slight sexual dimorphism for size in house sparrows. The mechanisms by which assortative mating occurs are in general little understood.

6.5 Conclusions

There has been considerable recent speculation about whether animals in natural populations mate with individuals having some optimum level of identical genes in common ('optimal outbreeding': see Bateson 1978, Shields 1982, 1983). Experimental evidence has been obtained that Japanese quail prefer to associate with individuals having an intermediate level of relatedness (Bateson 1978, 1982, 1983), implying that birds may use morphological cues to select a partner of the optimum type. Some evidence has been obtained here that house sparrows mate assortatively for size, though the result must be treated with caution as it was only found to be the case at one of the two study sites. This may imply that active mate choice occurs in house sparrows. It might be hypothesised that such choice allows the mating of birds with some optimum level of relatedness, but no evidence for any departure from random mating was obtained from the biochemical genetic analysis of mate pairs.

6.6 Summary

Six biochemical polymorphisms were used in a detailed analysis of the mating types of house sparrows in two breeding populations. No evidence was obtained for any departure from random mating amongst these polymorphisms. There was a significant tendency amongst the loci and samples for the inbreeding coefficients of the successful breeders to be negative.

Correlations between the sizes of mates were investigated, and significant assortative mating was found with respect to weight and

tail-length in the larger (Sutton Bonington) sample.

APPENDICES

Appendix 1

List of principal chemicals used, with details of suppliers.

Reagent	Supplier	Order No.
Adenosine	Sigma	A 9251
Agar (Bacteriological No 1)	Oxoid	-
Amido Black 10B (Naphthol Blue Black)	Sigma	N 3005
L-Amino acid oxidase	Sigma	A 5147
3-Amino-9-ethyl carbazole	Sigma	A 5754
Boric acid	BDH	27410
Citric acid	BDH	10081
DL-Dithiothreitol (DTT)	Sigma	D 0632
Fast Blue RR salt	Sigma	F 0500
α-D-Glucose-1-phosphate	Sigma	G 1259
D-Glucose-6-phosphate	Sigma	P 8391
D-Glucose 6-phosphate dehydrogenase	Sigma	G 7878
Heparin 5000 I.U./ml	Weddel Pharm. Ltd	-
DL-Isocitric acid (Na_3 salt)	Sigma	I 1252
L-Leucyl glycyl glycine	Sigma	L 9750
L-Leucyl-L-tyrosine	Sigma	L 0501
LiOH	BDH	29073
MTT	Sigma	M 2128
α-Naphthyl propionate	Sigma	N 0376
Nicotinamide-adenine dinucleotide phosphate (NADP)	BDH	42051
Nigrosin	Sigma	N 4754
Nitroso R salt	Hopkin & Williams	630400
Nucleoside phosphorylase	Sigma	N 3003
Peroxidase	Sigma	P 8250
6-Phosphogluconate (Na_3 salt)	Sigma	P 7877
Phenazine methosulphate (PMS)	Sigma	P 9625
NaH_2PO_4	BDH	10245
Na_2HPO_4	BDH	10249
Starch	Connaught Labs. Ltd	-
Triton	Sigma	T 6878
Tris(hydroxymethyl)aminomethane (TRIS)	Sigma	T 1378
Xanthine oxidase	Sigma	X 1875

Appendix 2

Electromorph and allele frequencies for age and sex classes in each year at each study site for each locus.

G_{HW} = goodness of fit G-test for agreement with Hardy-Weinberg ratios.

Appendix 2

Locus	Site	Year	Class	N	Electromorphs				Allele Frequencies			G_{HW}
					AA	AB	BB	BC	A	B	C	
6PGD	SB	1980	AdM	41	1	2	38	0	0.037	0.951		4.00
			AdF	42	0	3	39	0	0.036	0.964		0.11
			J	6	0	1	5	0	0.083	0.917		0.09
			N	54	0	4	50	0	0.037	0.963		0.15
		1981	AdM	100	0	5	95	0	0.025	0.975		0.13
			AdF	79	0	3	76	0	0.019	0.981		0.58
			J	14	0	0	14	0		1.000		0.00
			N	166	0	2	164	0	0.006	0.994		0.12
		1982	AdM	52	0	2	50	0	0.019	0.981		0.04
			AdF	42	0	1	41	0	0.012	0.988		0.12
			J	3	0	0	3	0		1.000		0.00
			N	217	1	10	206	0	0.028	0.972		2.18

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs				Allele Frequencies			G_{HW}
					AA	AB	BB	BC	A	B	C	
6PGD	BR	1980	AdM	47	0	1	46	0	0.011	0.989		0.01
			AdF	57	0	5	52	0	0.044	0.956		0.23
			J	33	0	3	30	0	0.045	0.945		0.14
			N	19	0	0	19	0		1.000		0.00
		1981	AdM	34	0	4	30	0	0.059	0.941		0.25
			AdF	34	0	3	31	0	0.044	0.956		0.14
			J	14	0	0	14	0		1.000		0.00
			N	120	0	6	113	1	0.025	0.971	0.004	0.21
		1982	AdM	25	0	1	24	0	0.020	0.980		0.02
			AdF	46	0	2	44	0	0.022	0.978		0.04
			J	17	0	0	17	0		1.000		0.00
			N	170	0	12	158	0	0.035	0.965		0.44

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs					Allele Frequencies				G_{HW}
					AB	BB	BC	BD	DD	A	B	C	D	
PEPD3	SB	1980	AdM	37	1	33	0	2	1	0.014	0.932		0.054	3.78
			AdF	37	1	32	0	3	1	0.014	0.919		0.068	2.70
			J	5	0	5	0	0	0		1.000			0.00
			N	48	0	41	0	7	0		0.927		0.073	0.55
		1981	AdM	91	0	79	0	11	1		0.929		0.071	0.56
			AdF	70	2	60	0	8	0	0.014	0.929		0.057	0.77
			J	10	0	10	0	0	0		1.00			0.00
			N	153	2	143	0	7	1	0.007	0.964		0.029	0.77
		1982	AdM	48	0	41	0	7	0		0.927		0.073	0.55
			AdF	35	0	32	0	3	0		0.957		0.043	0.13
			J	2	0	2	0	0	0		1.000			0.00
			N	191	3	171	0	17	0	0.008	0.948		0.045	1.10

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs					Allele Frequencies				G_{HW}
					AB	BB	BC	BD	DD	A	B	C	D	
PEPD3	BR	1980	AdM	40	1	35	0	4	0	0.013	0.938		0.050	0.33
			AdF	53	1	45	0	7	0	0.009	0.925		0.066	0.65
			J	31	0	29	0	2	0		0.968		0.032	0.07
			N	18	0	17	0	1	0		0.972		0.028	0.03
		1981	AdM	32	1	27	0	4	0	0.016	0.922		0.063	0.42
			AdF	32	2	25	1	4	0	0.031	0.891		0.063	0.86
			J	14	0	10	1	3	0		0.857	0.036	0.107	0.67
			N	108	6	92	3	6	1	0.028	0.921	0.014	0.037	3.49
		1982	AdM	23	1	21	0	1	0	0.022	0.957		0.022	0.09
			AdF	41	1	37	0	3	0	0.012	0.951		0.037	0.21
			J	16	0	15	0	1	0		0.969		0.031	0.03
			N	151	2	131	1	17	0	0.007	0.934	0.003	0.056	1.42

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs				Allele Frequencies			G_{HW}
					AA	AB	BB	BD	A	B	D	
PEPD2	SB	1980	AdM	41	1	3	37	0	0.061	0.939		2.79
			AdF	42	0	5	35	2	0.060	0.917	0.024	0.64
			J	6	0	1	5	0	0.083	0.917		0.09
			N	57	0	9	48	0	0.079	0.921		0.77
		1981	AdM	100	0	9	89	2	0.045	0.945	0.010	0.64
			AdF	79	0	9	70	0	0.057	0.943		0.54
			J	14	0	4	10	0	0.143	0.857		0.67
			N	166	0	13	152	1	0.039	0.958	0.003	0.62
		1982	AdM	52	0	4	48	0	0.038	0.962		0.16
			AdF	42	0	7	35	0	0.083	0.917		0.64
			J	3	0	1	2	0	0.167	0.833		0.20
			N	217	0	26	189	2	0.060	0.935	0.005	1.93

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs				Allele Frequencies			G_{HW}
					AA	AB	BB	BD	A	B	D	
PEPD2	BR	1980	AdM	47	0	7	40	0	0.075	0.926		0.56
			AdF	57	0	4	35	0	0.035	0.965		0.22
			J	33	0	2	31	0	0.030	0.970		0.06
			N	19	0	1	18	0	0.026	0.974		0.03
		1981	AdM	34	0	2	32	0	0.029	0.971		0.06
			AdF	34	0	2	32	0	0.029	0.971		0.06
			J	14	0	0	14	0		1.000		0.00
			N	121	0	13	108	0	0.054	0.946		3.49
		1982	AdM	25	0	2	23	0	0.040	0.960		0.08
			AdF	46	0	5	41	0	0.054	0.946		0.29
			J	17	0	1	16	0	0.059	0.941		0.03
			N	170	0	19	151	0	0.056	0.944		1.13

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs			Allele Frequencies		G_{HW}
					AA	AB	BB	A	B	
IDHC	SB	1980	AdM	41	25	13	3	0.768	0.232	0.47
			AdF	42	19	15	8	0.631	0.369	2.26
			J	5	2	2	1	0.600	0.400	0.14
			N	57	21	31	5	0.640	0.360	1.92
		1981	AdM	99	62	30	7	0.778	0.222	1.42
			AdF	78	42	31	5	0.737	0.263	0.05
			J	14	10	3	1	0.821	0.179	0.87
			N	165	103	55	7	0.791	0.209	0.01
		1982	AdM	52	26	21	5	0.702	0.298	0.06
			AdF	42	20	20	2	0.714	0.286	1.25
			J	3	1	1	1	0.500	0.500	0.34
			N	217	124	72	21	0.737	0.263	4.28†

† refer to text

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs AA	Electromorphs AB	Electromorphs BB	Allele Frequencies A	Allele Frequencies B	G_{HW}
IDHC	BR	1980	AdM	47	28	15	4	0.755	0.245	0.83
			AdF	57	28	26	3	0.719	0.281	1.00
			J	33	15	17	1	0.712	0.288	2.47
			N	19	10	9	0	0.763	0.237	2.84
		1981	AdM	34	10	21	3	0.603	0.297	2.97
			AdF	34	16	14	4	0.676	0.324	0.12
			J	14	5	8	1	0.643	0.357	0.88
			N	120	65	50	7	0.733	0.267	0.44
		1982	AdM	25	17	7	1	0.820	0.180	0.06
			AdF	46	18	19	9	0.598	0.402	0.91
			J	17	13	4	0	0.882	0.118	0.53
			N	170	75	63	32	0.626	0.374	7.31†

† refer to text.

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs						Allele Frequencies			G_{HW}
					AA	AB	AC	BB	BC	CC	A	B	C	
GP1	SB	1980	AdM	42	2	9	0	28	3	0	0.155	0.810	0.036	2.16
			AdF	44	0	11	0	31	2	0	0.125	0.852	0.023	2.26
			J	3	0	1	0	2	0	0	0.167	0.833		0.20
			N	57	1	7	0	43	6	0	0.079	0.868	0.053	2.24
		1981	AdM	98	5	23	2	63	4	1	0.179	0.781	0.041	4.93
			AdF	80	2	18	3	55	1	0	0.156	0.806	0.025	7.27
			J	14	0	6	0	8	0	0	0.214	0.786		1.66
			N	164	5	35	1	120	3	0	0.140	0.848	0.012	1.67
		1982	AdM	52	2	14	1	34	0	1	0.183	0.788	0.029	9.59*
			AdF	42	0	11	1	29	2	0	0.143	0.845	0.036	2.58
			J	3	0	1	0	2	0	0	0.167	0.833		0.20
			N	217	3	40	3	164	7	0	0.113	0.864	0.023	2.86

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs						Allele Frequencies			G_{HW}
					AA	AB	AC	BB	BC	CC	A	B	C	
GP1	BR	1980	AdM	46	3	12	0	30	1	0	0.196	0.793	0.011	1.59
			AdF	55	3	11	0	40	1	0	0.155	0.836	0.009	2.78
			J	33	6	10	1	16	0	0	0.348	0.636	0.015	5.09
			N	19	1	3	2	12	1	0	0.184	0.737	0.079	4.81
		1981	AdM	34	2	7	0	20	4	1	0.162	0.750	0.088	4.44
			AdF	34	2	8	1	19	4	0	0.191	0.735	0.074	1.07
			J	14	2	1	0	10	1	0	0.179	0.786	0.036	6.66
			N	121	10	88	5	59	9	0	0.260	0.682	0.058	2.39
		1982	AdM	25	1	10	0	13	1	0	0.240	0.740	0.020	0.86
			AdF	46	1	15	0	29	1	0	0.185	0.804	0.011	0.79
			J	17	2	7	1	5	2	0	0.353	0.559	0.088	0.34
			N	170	9	66	4	86	5	0	0.259	0.715	0.026	2.25

* $p < 0.05$

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs										Allele Frequencies					G_{HW}
					AD	BB	BC	BD	CC	CD	CE	DD	DE	EE	A	B	C	D	E	
PEPT	SB	1980	AdM	41	1	0	0	1	0	1	0	38	0	0	0.012	0.012	0.012	0.963		0.11
			AdF	42	0	0	0	2	0	1	0	39	0	0		0.024	0.012	0.964		0.11
			J	6	0	0	0	0	0	0	0	6	0	0				1.000		0.00
			N	57	0	0	0	3	0	2	0	52	0	0		0.026	0.018	0.956		0.29
		1981	AdM	100	0	0	0	8	0	6	0	85	1	0		0.004	0.007	0.925	0.001	1.22
			AdF	79	0	0	1	5	0	3	0	69	1	0		0.038	0.025	0.930	0.006	2.76
			J	14	0	0	0	0	0	1	0	13	0	0			0.036	0.964		0.04
			N	166	0	0	0	10	0	3	0	152	1	0		0.030	0.009	0.958	0.003	0.62
		1982	AdM	52	0	0	0	4	0	6	0	41	1	0		0.038	0.058	0.894	0.010	1.30
			AdF	42	1	0	0	0	0	1	0	40	0	0	0.012		0.012	0.976		0.05
			J	3	0	0	0	0	0	0	0	3	0	0				1.000		0.00
			N	217	5	0	0	6	0	11	0	194	1	0	0.012	0.014	0.025	0.947	0.002	1.29

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs										Allele Frequencies					G_{HW}
					AD	BB	BC	BD	CC	CD	CE	DD	DE	EE	A	B	C	D	E	
PEPT	BR	1980	AdM	47	0	1	0	3	0	4	0	36	3	0		0.053	0.043	0.872	0.032	4.17
			AdF	57	0	0	0	4	0	2	0	50	1	0		0.035	0.018	0.939	0.009	0.46
			J	33	0	0	0	3	0	2	0	26	1	1		0.045	0.030	0.879	0.045	5.77
			N	19	0	0	0	0	0	0	0	19	0	0				1.000		0.00
		1981	AdM	34	0	0	0	3	0	2	0	28	1	0		0.044	0.029	0.912	0.015	0.58
			AdF	34	0	1	0	1	0	3	0	29	0	0		0.044	0.044	0.912		5.46
			J	14	0	0	0	0	0	0	0	12	2	0				0.929	0.071	0.15
			N	121	0	0	0	11	0	5	1	102	2	0		0.045	0.025	0.917	0.012	5.20
		1982	AdM	25	0	0	0	1	0	0	0	21	3	0		0.020		0920	0.060	0.35
			AdF	46	0	0	1	4	0	3	0	36	2	0		0.054	0.043	0.880	0.022	2.52
			J	17	1	0	0	1	0	0	0	13	2	0	0.029	0.029		0.882	0.059	0.53
			N	170	0	0	0	10	1	7	1	145	6	0		0.029	0.029	0.921	0.021	5.82

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs							Allele Frequencies ± S.E.				
					A	AB	AC	B	BC	C	O	A	B	C	O	G_{HW}
EST2	SB	1980	AdM	42	1	5	0	36	0	0	0	0.076±0.029	0.867±0.070	-	0.056	0.71
			AdF	44	0	5	0	38	0	0	1	0.057±0.023	0.823±0.066	-	0.120	1.85
			J	4	0	0	0	3	1	0	0	-	-	-	-	-
			N	57	0	2	0	55	0	0	0	0.018±0.012	1.000±0.094	-	-0.018	0.00
		1981	AdM	98	2	14	1	77	2	1	1	0.090±0.019	0.785±0.042	0.021±0.010	0.105	1.50
			AdF	79	1	8	0	67	0	1	2	0.058±0.018	0.774±0.051	0.006±0.006	0.162	4.10
			J	14	0	2	0	12	0	0	0	0.072±0.049	1.001±0.189	-	-0.072	0.06
			N	166	3	14	0	146	2	1	0	0.053±0.012	0.865±0.035	0.009±0.995	0.072	5.12
		1982	AdM	52	4	2	0	45	1	0	0	0.064±0.025	0.768±0.066	0.010±0.010	0.158	9.16
			AdF	42	3	4	0	35	0	0	0	0.092±0.032	0.775±0.072	-	0.133	3.68
			J	3	0	0	0	3	0	0	0	-	-	-	-	-
			N	216	2	32	1	178	2	1	0	0.083±0.012	0.874±0.026	0.010±0.005	0.033	4.39

Appendix 2 (cont'd)

Locus	Site	Year	Class	N	Electromorphs							Allele Frequencies ± S.E.				
					A	AB	AC	B	BC	C	O	A	B	C	O	G_{HW}
EST2	BR	1980	AdM	48	1	6	0	40	1	0	0	0.078±0.027	0.869±0.063	0.070±0.010	0.043	0.70
			AdF	57	1	4	1	50	1	0	0	0.056±0.021	0.872±0.062	0.018±0.012	0.055	4.20
			J	33	0	3	0	30	0	0	0	0.046±0.026	1.000±0.123	-	-0.046	0.00
			N	19	0	4	0	15	0	0	0	0.105±0.050	1.001±0.162	-	-0.106	0.00
		1981	AdM	34	1	5	0	28	0	0	0	0.095±0.035	0.850±0.077	-	0.055	0.00
			AdF	34	0	5	0	28	1	0	0	0.073±0.032	0.999±0.122	0.015±0.015	-0.087	0.00
			J	14	0	1	0	12	1	0	0	0.036±0.035	1.004±0.190	0.036±0.035	-0.071	0.06
			N	120	2	13	0	102	3	0	0	0.066±0.016	0.882±0.040	0.012±0.007	0.039	2.16
		1982	AdM	25	0	5	0	20	0	0	0	0.100±0.035	1.00 ±0.077	-	-0.100	1.35
			AdF	46	3	3	0	39	0	0	1	0.070±0.026	0.731±0.068	-	0.199	2.33
			J	16	0	1	0	15	0	0	0	0.031±0.031	1.000±0.177	-	-0.032	0.06
			N	170	12	28	0	129	0	0	1	0.130±0.018	0.748±0.034	-	0.123	2.88

REFERENCES

ALLENDORF F.W., KNUDSEN K.L. & BLAKE G.M. 1982: Frequencies of null alleles at enzyme loci in natural populations of Ponderosa and red pine. Genetics **100**: 497-504.

ALLISON A.C. 1964: Polymorphism and natural selection in human populations. Cold Spr. Harb. Symp. Quant. Biol. **29**: 137-149.

ANDERSON T.R 1977: Reproductive responses of sparrows to a superabundant food supply. Condor **79**: 205-208.

ANDERSON T.R. 1978: Population studies of European sparrows in North America. Occ. Papers Mus. Nat. Hist. Univ. Kansas **70**: 1-58.

ANDERSON T.R. 1979: Experimental synchronisation of sparrow reproduction. Wilson Bull. **91**: 317-319.

ANDERSON W.W., 1974: Frequent multiple insemination in a natural population of Drosophila pseudoobscura. Amer. Natur. **108**: 709-711.

ANON 1977: (Office of Populations, Censuses and Surveys) Census 1971 England and Wales. Index of Place Names. 2 vols. HMSO, London.

ASHTON G.C. 1980: Mismatches in genetic markers in a large family study. Am. J. Hum. Genet. **32**: 601-613.

AVISE J.C. & SELANDER R.K. 1972: Evolutionary genetics of cave-dwelling fishes of the genus Astyanax. Evolution **26**: 1-19.

BACON P. 1979: Population Genetics of the Mute Swan. D.Phil. Thesis, Oxford University.

BAKER A.J. 1980: Morphometric differentiation in New Zealand populations of the house sparrow (Passer domesticus). Evolution **34**: 638-653.

BAKER C.M.A. 1965: Molecular Genetics of Avian Proteins. IV. The egg white proteins of the Golden Pheasant, Chrysolophus pictus L., and Lady Amherst's Pheasant, C. amherstiae Leadbeater, and their possible evolutionary significance. Comp. Biochem. Physiol. **16**: 93-101.

BAKER M.C. 1981: A muscle biopsy procedure in electrophoretic studies of birds. Auk **98**: 392-393.

BAKER M.C. 1983: The behavioural response of female Nuttall's white crowned sparrows to male song of natal and alien dialects. Behav. Ecol. Sociobiol. **12**: 309-315.

BAKER M.C., BAKER A.E.M., CUNNINGHAM M.A., THOMPSON D.B. & TOMBACK D.F. 1984: Reply to "Allozymes and song dialects: a reassessment". Evolution **38**: 449-451.

BAKER M.C. & FOX S.F. 1978: Dominance, survival and enzyme polymorphism in dark-eyed juncos, Junco hyemalis. Evolution **34**: 697-711.

BAKER M.C., SPITLER-NABORS K.J. & BRADLEY D.C. 1981: Early experience determines song dialect responsiveness of female sparrows. Science **214**: 819-821.

BAKER M.C., THOMPSON D.B., SHERMAN G.L., CUNNINGHAM M.A. & TOMBACK D.F. 1982 Allozyme frequencies in a linear series of song dialect populations. Evolution **36**: 1020-1029.

BANERJEE V. & BANERJEE M. 1970: Seasonal variations of erythrocyte number and haemoglobin content in a common Indian bird, Passer domesticus (Linnaeus). Zool. Anz. **185**: 31-35.

BATESON P. 1978: Sexual imprinting and optimum outbreeding. Nature **273**: 659-660.

BATESON P. 1982: Preferences for cousins in Japanese quail. Nature **295**: 286-237.

BATESON P. (Ed.) 1983: Mate Choice. Cambridge University Press.

BAVERSTOCK P.R., ADAMS M., POLKINGTHORNE R.W. & GELDER M. 1982: A sex-linked enzyme in birds - Z-chromosome conservation but no dosage compensation. Nature **296**: 736-766.

BEECHER M.D. & BEECHER I.M. 1979: Sociobiology of bank swallows: reproductive strategy of the male. Science **205**: 1282-1285.

BENNETT D. 1975: The T-locus of the mouse: a review. Cell **6**: 441-445.

BIRDSALL D.A. 1974: An analysis of selection at two loci in fluctuating populations of Microtus. Can. J. Zool. **52**: 1457-1462.

BIRDSALL D.A. & NASH D. 1973: Occurrence of successful multiple insemination of females in natural populations of deer mice, Peromyscus maniculatus. Evolution **27**: 106-110.

BIRKHEAD M.E. , BACON P.J. & WALTER P. 1983: Factors affecting the breeding success of the mute swan Cygnus olor. J. Anim. Ecol. **52**: 727-742.

BISHOP J.A., HARTLEY D.J. & PARTRIDGE G.G. 1977: The population dynamics of genetically determined resistance to warfarin in Rattus norvegicus from mid Wales. Heredity **39**: 389-398.

BOAG P.T. 1983: The heritability of external morphology in Darwin's ground finches, Geospiza, on Isla Daphne Major, Galapagos. Evolution **37**: 877-894.

BOAG P.T. & GRANT P.R. 1978: Heritability of external morphology in Darwin's finches. Nature **274**: 793-794.

BOROWSKY R. & KALLMAN K.D. 1976: Patterns of mating in natural populations of Xiphophorus (Pisces: Poeciliidae). I: X. maculatus from Belize and Mexico. Evolution **30**: 693-706.

BRACKBILL H. 1960: Determinate laying by house sparrows. Condor **62**: 499.

BRAY O.E., KENELLY J.J. & GUARINO J.L. 1975: Fertility of eggs produced on territories of vasectomised red-winged blackbirds. Wilson Bull. **87**: 187-195.

BROK-SIMONI F., ASHEKENAZI Y.E., RAMOT B., & HOLTZMAN F. 1976: The diurnal rhythm of enzymes in human red cells. Brit. J. Haem. **32**: 601.

BROUN M. 1972: Apparent migratory behaviour in the house sparrow. Auk **98**: 189-191.

BRUSH A.H. 1976: Waterfowl feather proteins: analysis of use in taxonomic studies. J. Zool. **179** 467.

BRUSH A.H. 1980: Patterns in the amino-acid composition of avian epidermal proteins. Auk **97**: 742-753.

BULATOVA N.S., RADJABU S.I., & PANOV E.N. 1972: Karyological description of three species of the genus Passer. Experientia **28**: 1369-1371.

BUMPUS H.C. 1898: The variations and mutations of the introduced sparrow. Biol. Lectures, Marine Biol. Lab., Woods Hole. 1-15.

BURNS J.T., CHENG K.M. & McKINNEY F.M. 1980: Forced copulation in captive mallards. I. Fertilization of eggs. Auk **97**: 875.

BUSH F.M. 1967: Developmental and population variation in electrophoretic properties of dehydrogenases, hydrolases and other blood proteins in the house sparrow Passer domesticus. Comp. Biochem. Physiol. **22**: 273-287.

BUSH F.M. & FARRAR W.W. 1969: Dissociation-reassociation of lactate dehydrogenase: Reversed isozyme migration and kinetic properties. Proc. Soc. Exp. Biol. Med. **131**: 13-15.

BUSH F.M., PRICE J.R. & TOWNSEND J.I. 1970: Plasma esterases, their definitions and status as isozymes in the house sparrow. Int. J. Biochem. **1**: 85-107.

CALHOUN J.B. 1947: The role of temperature and natural selection in relation to the variations in the size of the English sparrow in the United States. Amer. Natur. **81**: 203.

CASTROVIEJO J., CHRISTIAN L.C. & GROPP A. 1969: Karyotypes of 4 species of birds of the families Ploceidae and Paridae. J. Hered. **60**: 134-136.

CAVALLI-SFORZA L.L. & BODMER W.F. 1971: The Genetics of Human Populations. W.H. Freeman, San Francisco.

CHAKRABORTY R. & FERREL R.E. 1982: Efficient choice of genetic markers for paternity testing and correlation of paternity index and exclusion probability. Forensic Science Intl. **19**: 113-124.

CHAKRABORTY R. & RYMAN N. 1981: Use of odds of paternity computations in determing the reliability of single exclusions in paternity testing. Hum. Hered. **31**: 363-369.

CHEKE A.S. 1967: Notes on the ageing and sexing of juvenile house sparrows. Ringers' Bulletin **3**: 7-8.

CHEKE A.S. 1969: Mechanism and consequences of hybridization in sparrows Passer. Nature 222: 179-180.

CHEKE A.S. 1972: Movements and dispersal among house sparrows, Passer domesticus, at Oxford, England. In: KENDEIGH S.C. & PINOWSKI J. (Eds.) Productivity, Population Dynamics and Systematics of Granivorous Birds. Polish Scientific Publishers, Warsaw.

CHENG K.M., BURNS J.T. & McKINNEY F. 1983: Forced copulation in captive mallards: III. Sperm competition. Auk 100: 302-310.

CHITTY D. 1967: The natural selection of self-regulatory behaviour in animal populations. Proc. Ecol. Soc. Australia. 2: 51-78.

CHRISTIANSEN F.B. & FRYDENBERG O. 1976: Selection component analysis of natural polymorphisms using mother-offspring samples of successive cohorts. In: KARLIN S. & NEVO E.(Eds.) Population Genetics and Ecology pp207-301. Academic Press, New York.

CLARKE B. 1975: The contribution of ecological genetics to evolutionary theory: detecting the direct effects of natural selection on particular polymorphic loci. Genetics Suppl. 79: 101-113.

CLARKE B. & O'DONALD P. 1964: Frequency-dependent selection. Heredity 19: 201-206.

COCHRANE B.J. & RICHMOND R.C. 1979: Studies of Esterase-6 in Drosophila melanogaster I. The genetics of post-translational modification. Biochem. Genet. 17: 167-183.

COCKERHAM C.C. 1973: Analysis of gene frequencies of mates. Genetics 74: 701-712.

COCKERHAM C.C. & WEIR B.S. 1977: Digenic descent measures for finite populations. Genetical Research 30: 121-147.

COLE S.R. & PARKIN D.T. 1981: Enzyme polymorphism in the house sparrow, Passer domesticus. Biol. J. Linn. Soc. 15: 13-22.

CONAHAN M.A. 1970: The comparative accuracy of the likelihood ratio and X-squared as approximations to the exact multinomial test. Unpublished Ed. D. Dissertation, Lehigh University (Cited by Sokal & Rohlf 1981).

COOKE F., FINNEY G.H. & ROCKWELL R.F. 1976: Assortive mating in lesser snow geese, Anser caerulescens. Behav. Genet. 6: 127-140.

COOMBS C.J.F. 1960: Observations on the rook Corvus frugilegus in Southwest Cornwall. Ibis 102: 394-419.

CORNELIUS L.W. 1969: Field notes on Salmonella infection in Greenfinches and House sparrows. Bull. Wildlife Dis. Ass. 5: 142-142.

COTTERMAN C.W. 1941: Relatives and human genetic analysis. Scientific Monthly 53: 227-234.

COTTERMAN C.W. 1947: A weighting system for the estimation of gene frequencies from family records. Contr. Lab. Vertebr. Biol. Univ. Mich. 33: 1-21.

COTTERMAN C.W. 1954: Gene frequencies in nonexperimental populations. In: KEMPTHORNE O. et al. (Eds.). Statistics and Mathematics in Biology pp.449-465. Iowa State College Press, Iowa.

CROW J.F. 1981: The neutralist-selectionist controversy: an overview. In: HOOK E.B. (ed.) Population and Biological Aspects of Human Mutation. pp.3-14. Academic Press, New York.

CROW J.F. & KIMURA M. 1970: An Introduction to Population Genetics Theory. Harper & Row, New York.

CURIE-COHEN M. 1982: Estimates of inbreeding in a natural population: a comparison of sampling properties. Genetics 100: 339-358.

DARWIN C. 1859: On the Origin of Species by Means of Natural Selection. John Murray, London.

DAWSON D.G. 1970: Estimation of grain loss due to sparrows (Passer domesticus) in New Zealand. N. Zealand J. Agric. Res. 13: 681-688.

DAWSON D.G. 1972: The Breeding Ecology of House Sparrows. D. Phil. Thesis, Oxford University.

DAWSON D.G. & BULL P.C. 1970: A questionnaire survey of bird damage to fruit. N. Zealand J. Agric. Res. 13: 326-371.

DAY T.H., HILLIER P.C. & CLARKE B.C. 1974: The properties of genetically polymorphic isozymes of alcohol dehydrogenase in D. melanogaster. Biochem. Genet. **11** : 141-153.

DE BENEDICTUS P.A. 1978: Are populations characterised by their genes or by their genotypes? Amer. Natur. **112**: 155-175.

DE CAMINOS S.J., BUCHER E.H. & BLANCO A. 1981: Temporal variations of allele frequencies in the eared dove (Zenaida auriculata). Biochem. Genet. **19**: 1163-1167.

DIXON W.S, BROWN M.B., ENGELMAN L., FRANE J.W., HILL M.A., JENNRICH R.I. & TOPOREK J.D. 1983: BMDP Statistical Software. New edition. University of California Press, Berkeley.

DOBZHANSKY T. 1955: A review of some fundamental concepts and problems of population genetics. Cold Spring Harbor Symp. Quart. Biol. **20**: 1-15.

DUVALL S.W., BERNSTEIN I.S. & GORDON T.P. 1976: Paternity and status in a rhesus monkey group. J. Reprod. Fertil. **47**: 25-31.

DYER M.I., PINOWSKI J. & PINOWSKA B. 1977: Population dynamics. In: PINOWSKI J. & KENDEIGH S.C. Granivorous Birds in Ecosystems. pp.53-106. Cambridge University Press.

ELANDT-JOHNSON R.C. 1971: Probability Models and Statistical Methods in Genetics. Wiley, New York.

EMLEN J.T. 1954: Territory, nest-building and pair formation in the cliff swallow. Auk **71**: 16-35.

ERICKSON C.J. & ZENONE P.G. 1976: Courtship differences in male ring doves: avoidance of cuckoldry? Science **192**: 1353-1354.

ERICKSON C.J. & ZENONE P.G. 1978: Aggressive courtship as a means of avoiding cuckoldry. Anim. Behav. **26**: 307.

EVANS P.G.H. 1980: Population Genetics of the European Starling, Sturnus vulgaris. D.Phil. Thesis, Oxford University.

EVERITT B.S. 1977: The Analysis of Contingency Tables. Chapman & Hall, London.

FALCONER D.S. 1981: Introduction to Quantitative Genetics. 2nd Edition. Longman, London.

FERGUSON A. 1980: Biochemical Systematics and Evolution. Blackie, Glasgow & London.

FINDLAY C.S. & COOKE F. 1982: Breeding synchrony in the lesser snow goose, Anser caerulescens caerulescens. I. Genetic and environmental components of hatch date variability and their effects on hatch synchrony. Evolution **36**: 342-351.

FINDLAY C.S. & COOKE F. 1983: Genetic and environmental components of clutch size variance in a wild population of lesser snow geese, Anser caerulescens caerulescens. Evolution **37**: 724-734.

FITCH H.S. 1949: Sparrow adopts kingbirds. Auk. **66**: 368-369.

FLEISCHER R.C. 1983: A comparison of theoretical and electrophoretic assessments of genetic structure in populations of the house sparrow, Passer domesticus. Evolution **37**: 1001-1009.

FLEISCHER R.C. & JOHNSTON R.F. 1982: Natural selection of body size and proportions in house sparrows. Nature **298**: 747-749.

FLEISCHER R.C., JOHNSTON R.F. & KLITZ W.J. 1983: Allozymic heterozygosity and morphological variation in house sparrows. Nature **304**: 628-630.

FLUX J.E.C. & FLUX M.M. 1982: Artificial selection and gene flow in wild starlings, Sturnus vulgaris. Naturwissenschaften **69**: 96-97.

FOLTZ D.W. 1981a: Genetic evidence for long-term monogamy in a small rodent, Peromyscus polionatus. Amer. Natur. **117**: 665-675.

FOLTZ D.W. 1981b: Genetic measures of inbreeding in Peromyscus. J. Mamm. **62**: 470-476.

FOLTZ D.W. & HOOGLAND J.L. 1981: Analysis of the mating system in the black-tailed prairie dog, Cynomys ludovicianus by likelihood of paternity. J. Mamm. **62**: 706-712.

FORD E.B. 1975: Ecological Genetics. 4th edn. Chapman & Hall, London.

FOWLE K.E., CLINE J.H., KLOSTERMAN E.W. & PARKER C.F. 1967: Transferrin genotypes and their relationship with blood constituents, fertility and cow productivity. J. Anim. Sci. **28**: 1226-1231.

FRANKLIN L. & LEWONTIN R.C. 1970: Is the gene the unit of selection? Genetics **65**: 707-734.

FRELINGER J.A. 1971: Maternally derived transferrin in pigeon squabs. Science **171**: 1260-1261.

FRELINGER J.A. 1972: The maintenance of transferrin polymorphism in pigeons. Proc. Natl. Acad. Sci. U.S.A. **69**: 326-329.

FRELINGER J.A. 1973: Chemical basis of transferrin polymorphism in pigeons. Anim. Blood Grps. Bochem. Genet. **4**: 35-40.

FRELINGER J.A. & CROW J.F. 1973: Transferrin polymorphism and the Hardy Weinberg ratios. Amer. Natur. **107**: 314-317.

FUJIOKA M. & YAMAGISHI S. 1981: Extramarital and pair copulations in the cattle egret. Auk **98**: 134-144.

GAHNE B. 1966: Studies of the inheritance of electrophoretic forms of transferrin, albumin, pre-albumin and plasma esterases of horses. Genetics **53**: 681-694.

GAINES M.S. 1981: Importance of genetics to population dynamics. In: SMITH M.H. & JOULE J. 1981: Mammalian Population Genetics. University of Georgia, Athens, Georgia.

GARNETT M.C. 1981: Body size, its heritability and influence on juvenile survival among great tits, Parus major. Ibis **123**: 31-41.

GARTEN C.T. 1977: Relationships between exploratory behaviour and genic heterozygosity in the oldfield mouse. Anim. Behav. **25**: 328-332.

GILBERT D.G., RICHMOND R.C. & SHEEHAN K.B. 1982: Studies of esterase-6 in Drosophila melanogaster. V. Progeny production and sperm use in females inseminated by males having active or null alleles. Evolution **35**: 21-37.

GILLESPIE J.H. & LANGLEY C.H. 1974: A general model to account for enzyme variation in natural populations. Genetics **76**: 837-848.

GLADSTONE D.E. 1979: Promiscuity in monogamous colonial birds. Amer. Natur. **114**: 545-557.

GRANT P.R. 1972: Centripetal selection and the house sparrow. Syst. Zool. **21**: 23-30.

GRANT V. 1977: Organismic Evolution W.H. Freeman, San Francisco.

GUNDEL H. & REETZ I. 1981: Exclusion probabilities obtainable by biochemical polymorphisms in dogs. Anim. Blood Groups and Biochem. Genet. **12**: 123-132.

GUSTAFSON D.P. & MOSES H.E. 1953: The English Sparrow as a natural carrier of Newcastle disease virus. Am. J. Vet. Res. **14**: 581-585.

HABER M. 1980: Detection of inbreeding effects by the χ^2 test on genotypic and phenotypic frequencies. Am. J. Hum. Genet. **32**: 754-760.

HALLIDAY T.R. 1983: The study of mate choice. In: BATESON P. Mate Choice pp.3-32. Cambridge University Press.

HAMILTON G.D. 1952: English sparrow feeding young Eastern kingbirds. Condor **54**: 316.

HANKEN J. & SHERMAN P.W. 1981: Multiple paternity in Belding's ground squirrel litters. Science **212**: 351-353.

HARRIS H. 1966: Enzyme polymorphisms in Man. Proc. Roy. Soc. B. **164**: 298-310.

HARRIS H.G. & HOPKINSON D.A. 1972: Average heterozygosity per locus in man: an estimate based on the incidence of enzyme polymorphism. Ann. Hum. Genet. Lond. **36**: 9-20.

HARRIS H.G. & HOPKINSON D.A. 1978: Handbook of Enzyme Electrophoresis in Human Genetics. North-Holland, Amsterdam.

HARRIS J.A. 1911: A neglected paper on natural selection in the English sparrow. Amer. Natur. **45**: 314-318.

HARRISON J.M. 1961: The significance of some plumage phases of the house sparrow, Passer domesticus (Linnaeus) and the Spanish sparrow, Passer hispaniolensis (Temminck). Bull. Brit. Ornithologists' Club **81**: 96-101.

HARTL D. 1977: How does the genome congeal? In: LEVIN S. (Ed.) Measuring Selection in Natural Populations. Lecture Notes in Biomathematics **19**, Springer-Verlag, Berlin.

HEDRICK P.W., GINEVAN M.E. & EWING E.P. 1976: Genetic polymorphism in heterogeneous environments. Ann. Rev. Ecol. Syst. **7**: 1-32.

HENDERSON B.A. 1976: Protein polymorphism in the egg albumen of red grouse (L. lagopus scoticus). Heredity **36**: 413-415.

HENDERSON B.A. 1977: The genetics and demography of a high and low density of red grouse. J. Anim. Ecol. **46**: 581-592.

HICKEY W.A. & CRAIG G.B. Jr. 1966: Genetic distortion of sex ratio in a mosquito, Aedes aegypti. Genetics **53**: 1177-1196.

HILL W.G. 1974: Estimation of linkage disequilibrium in randomly mating populations. Heredity **33**: 229-239.

HIRSCH J. & VETTA A. 1978: Gli errori concettuali dell'analisi genetico-comportamentale. Ricerche di Psicologia **2**: 47-67.

HODSON N.L. & SNOW D.W. 1965: The Road Deaths Inquiry. Bird Study **12**: 90-99.

HOOGLAND J.L. & SHERMAN P.W. 1978: Advantages and disadvantages of bank swallow, Riparia riparia, coloniality. Ecol. Monogr. **46**: 33-58.

HUBALEK Z. 1977: Spread of fungi and other microorganisms by sparrows: a review. Intl. Stud. Sparrows. **10**: 7-25.

HUBBY J.I. & LEWONTIN R.G. 1966: A molecular approach to the study of heterozygosity in natural populations. I. The number of alleles at different loci in Drosophila pseudoobscura. Genetics **54**: 577-594.

JAMES F.C. 1983: Environmental components of morphological differentiation in birds. Science **221**: 184-186.

JAMIESON A. 1965: The genetics of transferrins in cattle. Heredity **20**: 419-441.

JOHNSTON R.F. 1969a: Character variation and adaptation in European sparrows. Syst. Zool. **18**: 206-231.

JOHNSTON R.F. 1969b: Taxonomy of house sparrows and their allies in the Mediterranean basin. Condor **71**: 129-139.

JOHNSTON R.F. 1972: Color variation and natural selection in Italian sparrows. Esteatto da Bolletino de Zoologia **39**: 351-362.

JOHNSTON R.F. 1973: Evolution in the house sparrow: IV. Replicate studies in phenetic covariation. Syst. Zool. **22**: 219-226.

JOHNSTON R.F. & FLEISCHER R.C. 1981: Overwinter mortality and sexual size dimorphism in the house sparrow. Auk **98**: 503-511.

JOHNSTON R.F. & KLITZ W.J. 1977: Variation and evolution in a granivorous bird: the house sparrow. In: PINOWSKI J. & KENDEIGH S.C. (Eds.) Granivorous birds in ecosystems pp.15-51. International Biological Programme **12**, Cambridge University Press.

JOHNSTON R.F., NILES D.M. & ROHWER S.A. 1972: Hermon Bumpus and natural selection in the house sparrow, Passer domesticus. Evolution **28**: 20-31.

JOHNSTON R.F. & SELANDER R.K. 1964: House sparrows: Rapid evolution of races in North America. Science **144**: 548-550.

JOHNSTON R.F. & SELANDER R.K. 1971: Evolution in the house sparrow: II. Adaptive differentiation in North American populations. Evolution **25**: 1-28.

JOHNSTON R.F. & SELANDER R.K. 1972: Variation, adaptation and evolution in the North American house sparrow. In: KENDEIGH S.C. & PINOWSKI J. (Eds.) Productivity, Population Dynamics and Systematics of Granivorous Birds. Polish Scientific Publishers, Warsaw.

JOHNSTON R.F. & SELANDER R.K. 1973: Evolution in the house sparrow: II. Variation in size and sexual dimorphism in Europe and North and South America. Amer. Natur.**107**: 373-390.

KALMBACH E.R. 1940: Economic status of the English sparrow in the United States. U.S. Dept. Agric. Tech. Bull. **711**: 66.

KAN Y.W. & DOZY A.M. 1980: Evolution of the hemoglobin S and C genes in world populations. Science **209**: 388-391.

KARLSON J. & NILSSON S.G. 1977: The influence of nest-box area on clutch-size in some hole-nesting passerines. Ibis **119**: 207-211.

KAT P.W. 1982: The relationship between heterozygosity for enzyme loci and developmental homeostasis in peripheral populations of aquatic bivalves, Unionidae. Amer. Natur. **119**: 824-832.

KELLER M.E. 1979: Breeding behavior and reproductive success of Chipping sparrows in northwestern Minnesota. M.S. Thesis, University of N. Dakota.

KENDEIGH S.C. (Ed.) 1973: A symposium on the house sparrow, Passer domesticus, and European tree sparrow, P. montanus, in North America. Ornithol. Monogr. **14**, American Ornithologists' Union.

KENDEIGH S.C., DOL'NIK V.R. & GAVRILOV V.M. 1977: Avian Energetics. In: PINOWSKI J. & KENDEIGH S.C. (Eds) Granivorous Birds in Ecosystems. pp129-197. Cambridge University Press.

KENDEIGH S.C. & PINOWSKI J. (Eds.) 1972: Productivity, Population Dynamics and Systematics of Granivorous Birds. Polish Scientific Publishers, Warsaw.

KIMURA M. 1983: The Neutral Theory of Molecular Evolution. Cambridge University Press.

KIMURA M. & CROW J.F. 1964: The number of alleles that can be maintained in a finite population. Genetics **49**: 725-738.

KING J.L. & JUKES T.H. 1969: Non-Darwinian evolution: random fixation of selectively neutral mutations. Science **164**: 788-798.

KLEIMAN D.G. 1977: Monogamy in mammals. Q. Rev. Biol. **52**: 39-69.

KLITZ W. 1972: Genetic Consequences of Colonisation with Subsequent Expansion: The House Sparrow in North America. Ph.D.Thesis (part), University of Kansas.

KLITZ W. 1973: Empirical population genetics of the North American house sparrow. Ornithological Monographs **14**: 39-48.

KLOMP H. 1970: The determination of clutch-size in birds. A review. Ardea **58**: 1-124.

KNOX A.G. 1980: Feather protein as a source of avian taxonomic information. Comp. Biochem. Physiol. **65B**: 45-54.

KOBLIANSKY E. & LIVSHITS G. 1983: Relationship between levels of biochemical heterogeneity and morphological variability in human populations. Am. J. Hum. Genet. **47**: 215-223.

KOEHN R.K., ZERA A.J. & HALL J.G. 1983: Enzyme polymorphism and natural selection. In: NEI M. & KOEHN R.K. (Eds.) Evolution of Genes and Proteins pp115-136. Sinauer, Sunderland, Ma.

KOENIG W.D. & PITELKA F.A. 1979: Relatedness and inbreeding: counterploys in the communally-nesting acorn woodpecker. Science **206**: 1103-1105.

KOJIMA K. & YARBROUGH K.M. 1967: Frequency-dependent selection at the esterase-6 locus in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA **57**: 645-649.

KOSTELECKA-MYRCHA A., PINOWSKI J. & TOMEK T. 1971: Changes of the respiratory function of the blood of nestling common sparrows, Passer domesticus L. during their development. Bull. Acad. Pol. Sci. Ser. Sci. Biol. **19**: 639-645.

KOVÁCH A.G.B., SZÁSZ E. & PILAYER N. 1969: The mortality of various avian and mammalian species following blood loss. Acta. Physiol. Acad. Sci. Hung. **35**: 109-126.

KREYMBORG A. 1911: Ueber das sich-tot-legen-lassen von Vögeln. Orn. Monatsschr. **36**: 86-88.

KUSHLAN J.A. 1973: Promiscuous mating behaviour in the white ibis. Wilson Bull. **85**: 331-332.

LACK D. 1947: The significance of clutch size. Ibis **89**: 302-352.

LACK D. 1968: Ecological Adaptations for Breeding in Birds. Methuen, London.

LANDE R. & ARNOLD S.J. 1983: The measurement of selection on correlated characters. Evolution **37**: 1210-1226.

LANDSTEINER K., LONGSWORTH L.G. & VAN DER SCHEER J. 1938: Electrophoresis experiments with egg albumins and hemoglobins. Science **88**: 83-85.

LANGLEY C.H. & CROW J.F. 1974: The direction of linkage disequilibrium. Genetics **78**: 937-941.

LANGLEY C.H., VOELKER R.A., LEIGH BROWN A.J. et al. 1981: Null allele frequencies at allozymic loci in natural populations of Drosophila melanogaster. Genetics **99**: 151-156.

LANGLEY C.H., TOBARI Y.N. & KOJIMA K.-I. 1974: Linkage disequilibrium in natural populations of Drosophila melanogaster. Genetics **78**: 921-936.

LATHROP G.M., HOOPER A.B., HUNTSMAN J.W. & WARD R.H. 1983: Evaluating pedigree data. I. The estimation of pedigree error in the presence of marker mistyping. Am. J. Hum. Genet. **35**: 241-262.

LEARY R.F., ALLENDORF F.W. & KNUDSON K.L. 1983: Developmental stability and enzyme heterozygosity in rainbow trout. Nature **301**: 71-72.

LEIGH BROWN A.J., & LANGLEY C.H. 1979b: Reevaluation of the level of genic heterozygosity in natural populations of D. melanogaster by two-dimensional electrophoresis. Proc. Natl. Acad. Sci. **76**: 2381-2384.

LEVENE H. 1949: On a matching problem arising in genetics. Annals of Mathematical Statistics **20**: 91-94.

LEVINE L. 1982: Observed frequencies of multiply inseminated Drosophila pseudoobscura females in natural populations: a note of caution and explanation. Amer. Natur. **120**: 420-421.

LEWONTIN R.C. 1974: The Genetic Basis of Evolutionary Change. Columbia University Press, New York.

LEWONTIN R.C., GINZBURG L.R. & TULJAPURKAR S.D. 1978: Heterosis as an explanation for large amounts of genic polymorphism. Genetics **88**: 149-170.

LI C.C. 1955: Population genetics. University of Chicago Press, Chicago.

LIMPKIN S., KESSEL K., ZENONE P. & ERICKSON C.J. 1982: Proximity between the sexes in ring doves: social bonds or surveillance? Anim. Behav. **30**: 506-513.

LKB 1977: Analytical electrofocusing in thin layer polyacrylamide gels. Application Note 250, LKB, Sweden.

LONG J.L. 1981: Introduced Birds of the World. David and Charles, Newton Abbot.

LORD, R.D. et al. 1974: Assessment of bird involvement in three urban St. Louis encphalitis epidemics. Am. J. Epidemiol. **99**: 364-367.

LOUKAS M., KRIMBAS C.B. & MORGAN K. 1980: The genetics of Drosophila subobscura populations. XIV. Further data on linkage disequilibria. Genetics **95**: 757-768.

LOUKAS M., KRIMBAS C.B. & VERGINI Y. 1979: The genetics of Drosophila subobscura populations. IX. Studies on linkage disequilibrium in four natural populations. Genetics **93**: 497-523.

LOWTHER J.K. 1961: Polymorphism in the white-throated sparrow, Zonotrichia albicollis (Gmelin). Can. J. Zool. **39**: 281-292.

LOWTHER P.E. 1977: Selection intensity in North American house sparrows, Passer domesticus. Evolution **31**: 649-656.

LOWTHER P.E. 1979a: The nesting biology of House Sparrows in Kansas. Kansas Ornithol. Soc. Bull. **30**: 23-28.

LOWTHER P.E. 1979b: Growth and dispersal of nestling House Sparrows: sexual differences. Inland Bird Banding **51**: 23-29.

LOWTHER P.E. 1979c: Overlap of house sparrow broods in the same nest. Bird Banding **50**: 160-162.

LUCOTTE G. & KAMINSKI M. 1976: Molecular heterosis at the conalbumin locus in the ring-necked pheasant (Phasianus colchicus). Theoretical and Applied Genetics **48**: 251-253.

MACDONALD J.W. & BROWN D.D. 1974: Salmonella infection in wild birds in Britain. Vet. Rev. **94**:321-322.

MACKOWICZ R., PINOWSKI J. & WIELOCH M. 1970: Biomass production by house sparrow (Passer domesticusL.) and tree sparrow (Passer montanus L.) populations in Poland. Ecol. Polska. **A 18**: 465-501.

MACROBERTS M.H. 1973: Extramarital courting in lesser black-backed and herring gulls. Zeitschrift für Tierpsychologie **32**: 62-74.

MAJUMDER P.P & NEI M. 1983: A note on positive identification of paternity by using genetic markers. Hum. Hered. **33**: 29-35.

MANWELL C. & BAKER C.M.A. 1975: Molecular genetics of avian proteins. XIII. Protein polymorphism in three species of Australian passerines. Aust. J. Biol. Sci. **28**:545-557.

MAYNARD SMITH S., PENROSE L.S. & SMITH C.A.B. 1961: Mathematical Tables for Research Workers in Human Genetics. Churchill, London.

MAYR E. 1963: Animal Species and Evolution. Harvard University Press, Cambridge, Mass.

McANDREW B.J., WARD R.D. & BEARDMORE J.A. 1982: Lack of relationship between morphological variance and enzyme heterozygosity in the plaice, Pleuronectes platessa. Heredity **48**: 117-126.

McCLURE H. & CEDENO R. 1955: Techniques for taking blood samples from living birds. J. Wildlife Mgment. **19**: 477-478.

McCRACKEN G.F. & BRADBURY J.W. 1977: Paternity and genetic heterogeneity in the polygynous bat, Phyllostomus herstatus. Science **198**: 303-306.

McGILLIVRAY W.B. 1980: Nest grouping and productivity in the house sparrow. Auk **97**: 396-399.

McGILLIVRAY W.B. 1983: Intraseasonal reproductive costs for the House Sparrow, Passer domesticus. Auk **100**: 25-32.

McLELLAN T., FERRO-LUZZI AMES G. & NIKAIDO K. 1983: Genetic variation in protein: comparison of one-dimensional and two-dimensional gel electroporesis. Genetics **104**: 381-390.

MEANLEY B. 1955: A nesting study of the little blue heron in eastern Arkansas. Wilson Bull. **67**: 84-99.

MERRITT R.B. & WU B.J. 1975: On the quantification of promiscuity (or 'Promiscuous maniculatus'?) Evolution **29**: 575-578.

METCALF R.A. & WHITT G.S. 1977: Intra-nest relatedness in the social wasp, Polistes metricus. Behav. Ecol. Sociobiol. **2**: 339-351.

MILKMAN R. & ZEITLER R.R. 1974: Concurrent multiple paternity in natural and laboratory populations of Drosophila melanogaster. Genetics **78**: 1191-1193.

MINEAU P. & COOKE F. 1979: Rape in the lesser snow goose. Behaviour **70**: 280-291.

MITCHELL C.J. & HAYES R.O. 1973: Breeding house sparrows, Passer domesticus, in captivity. Ornithol. Monogr. **14**: 39-48.

MITCHELL C.J., HAYES R.O., HOLDEN P. & HUGHES T.B. 1973: Nesting activity of the House Sparrow in Hale County, Texas, during 1968. Ornithol. Monogr. **14**: 49-59.

MOEED A. & DAWSON D.G. 1979: Breeding of starlings, Sturnus vulgaris, in nest boxes of various types. New Zealand J. Zool. **6**: 613-618.

MONTAG D.G. & DAHLGREN R.B. 1973: Transferrin polymorphism in the ring-necked pheasant. Auk **90**: 318-323.

MORTON N.E. 1955: Sequential tests for the detection of linkage. Am. J. Hum. Genet. **7**: 277-318.

MORTON N.E. 1957: Further scoring types in sequential linkage tests, with a critical review of autosomal and partial sex linkage in man. Am. J. Hum. Genet. **9**: 55-75.

MUELLER J.O., SMITHIES O. & IRWIN M.R. 1962: Transferrin variation in Columbidae. Genetics **47**: 1385-1392.

MULLER H.J. 1962: Studies in Genetics. Indiana University Press.

MURPHY E.C. 1978a: Breeding ecology of house sparrows: spatial variation. Condor **80**: 180-193.

MURPHY E.C. 1978b: Seasonal variation in reproductive output of house sparrows: the determination of clutch size. Ecology **59**: 1189-1199.

MURRAY J.J. 1964: Multiple mating and effective population size in Cepaea nemoralis. Evolution **80**: 283-291.

NAIK R.M. & MISTRY L. 1970: Breeding season and reproductive rate of Passer domesticus in Baroda, India. Intl. Stud. Sparrows **4**: 15-16.

NEI M. 1972: Genetic distance between populations. Amer. Natur. **106**: 283-292.

NEI M. 1975: Molecular Population Genetics and Evolution. North-Holland, Amsterdam.

NEI M. & ROYCHOUDHURY A.K. 1974: Sampling variances of heterozygosity and genetic distance. Genetics **76**: 379-390.

NERO R.W. 1951: Pattern and rate of cranial ossification in the house sparrow. Wilson Bull. **63**: 84-88.

NEWTON I. 1968: The moulting seasons of some finches and buntings. Bird Study **15**: 84-92.

NIE N.H., HULL C.H., JENKINS J.G., STEINBRENNER K. & BENT D.H. 1975: SPSS Statistical Package for the Social Sciences. 2nd Edition. McGraw-Hill, New York.

NIELSEN P.B. 1970: Theoretical exclusion percentages obtainable by blood grouping in pigs. Animal Blood Groups and Biochemical Genetics **1** : 53-62.

NIELSEN J.T. 1977: Polymorphism for the number of genes coding for salivary amylase in the bank vole Clethrionymus glareola. In: LEVIN S. (Ed.) Measuring Selection in Natural Populations pp.275-290. Lecture Notes in Biomathematics **19**, Springer-Verlag, Berlin.

NILES D.M. 1973: Geographic and seasonal variation in the occurrence of incompletely pneumatized skulls in the house sparrow. Condor **75**: 352-356.

NORTH C.A. 1968: A study of house sparrow populations and their movements in the vicinity of Stillwater, Oklahoma. Ph.D. Thesis, Oklahoma State University.

NOTTEBOHM F. & SELANDER R.K. 1972: Vocal dialects and gene frequencies in the chingolo sparrow, Zonotrichia capensis. Condor **74**: 137-143.

NOVOTNÝ I. 1970: Breeding bionomics growth and development of young house sparrows, Passer domesticus, Linné 1785. Acta. Sci. Nat. Acad. Sci. Bohemoslande Brno. **4**: 1-57.

OAKESHOTT J.G., GIBSON J.B., ANDERSON P.R., KNIBB W.R. & CHAMBERS G.K. 1982: Alcohol dehydrogenase and glycerol-3-phosphate clines in Drosophila melanogaster on different continents. Evolution **36**: 86-96.

O'BRIEN S.J., MITCHELL M.G. & LEVIN D.L. 1980: Correlative genetic variation in natural populations of cats, mice and man. Nature **288**: 580-583.

O'DONALD P. 1973: A further analysis of Bumpus' data: the intensity of natural selection. Evolution **27**: 398-404.

O'FARRELL P.M. 1975: High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. **250**: 4007-4021.

OGDEN A.L., MORTON J.R., GILMOUR D.G. & McDERMID E.M. 1962: Inherited variants in the transferrins and conalbumins of the chicken. Nature **195**: 1026-1028.

ORING L.W. 1982: Avian mating systems. In: FARNER D.S., KING J.R. & PARKER K.C. (Eds.), Avian Biology Vol.VI pp.1-92. Academic Press, New York.

ORMIAN M. 1979: [The application of the polymorphism of blood proteins and enzymes in parentage control] Pracemater Zootech. **18**: 17-25.

PACKARD G.C. 1967: House sparrows: Evolution of populations from the Great Plains and Colorado Rockies. Syst. Zool. **16**: 73-89.

PAGE R.E. & METCALF R.A. 1982: Multiple mating, sperm utilization and social evolution. Amer. Natur. **119**: 236-287.

PARKER G.A. 1970: Sperm competition and its evolutionary consequences in the insects. Biol. Rev. **45**: 525-568.

PARKIN D.T. 1984: The evolutionary genetics of the house sparrow, Passer domesticus. (Under review).

PARKIN D.T. & COLE S.R. 1984a: Genetic variation in the House Sparrow (Passer domesticus) in the east midlands of England. Biol. J. Linn. Soc., (in press).

PARKIN D.T. & COLE S.R. 1984b: Genetic differentiation and rates of evolution in some populations of the House Sparrow (Passer domesticus) in Australia and New Zealand. Heredity, (in press).

PEARSE T. 1940: Polygamy in the English sparrow. Condor **42**: 124-125.

PERRINS C.M. & JONES P.J. 1974: The inheritance of clutch size in the great tit, Parus major L. Condor **76**: 225-228.

PERRINS C.M. & MOSS D. 1975: Reproductive rates in the Great Tit Parus majorL. J. Anim. Ecol. **44**: 695-706.

PHILIPP E.E. 1973: Discussion remarks In: Law and Ethics of AID and Embryo Transfer pp.63-66. Ciba Found. Symp. New Series **17**.

PINOWSKI J. & KENDEIGH S.C. (Eds.) 1977: Granivorous Birds in Ecosystems. Their Evolution, Populations, Energetics, Adaptations, Impact and Control. International Biological Programme **12**, Cambridge University Press.

PINOWSKI J. & MYRCHA A. 1977: Biomass and production rates. In: PINOWSKI J. & KENDEIGH S.C. (Eds.) Granivorous Birds in Ecosystems. pp.107-126. Cambridge University Press.

PRICE T., MILLINGTON S. & GRANT R. 1983: Helping at the nest in Darwin's Finches as misdirected parental care. Auk **100**: 192-194.

PRUZAN-HOTCHKISS A., EJIANNE D. & FARO S.H. 1981: Sperm utilization in once-and-twice-marked Drosophila pseudoobscura females. Amer. Natur. **118**: 37-45.

PUHLMANN E. 1914: Das sich-tot-legen-lassen von Vögeln. Orn. Monatsschr **39**: 512-515.

RACINE R.R & LANGLEY C.H. 1980: Genetic heterozygosity in a natural population of Mus musculus assessed using two-dimensional electrophoresis. Nature **283**: 855-857.

REDFIELD J.A. 1973: The use of incomplete family data in the analysis of genetics and selection at the Ng locus in blue grouse, Dendragapus obscurus. Heredity **31**: 35-42.

REDFIELD J.A. 1974: Genetics and selection at the Ng locus in blue grouse, Dendragapus obscurus. Heredity **33**: 69-78.

REYNOLDS J., WEIR B.S. & COCKERHAM C.C. 1983: Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics **105**: 767-779.

RISING J.D. 1972: Age and seasonal variation in dimensions of house sparrows, Passer domesticus (L.), from a single population in Kansas. In: KENDEIGH S.C. & PINOWSKI J. (Eds.) Productivity, Population Dynamics and Systematics of Granivorous Birds. pp.327-336. Polish Scientific Publishers, Warsaw.

RISSMAN E.F. 1983: Detection of cuckoldry in ring doves. Anim. Behav. 31: 449-456.

ROBERTS D.F. 1977: Assortative matings in man: husband/wife correlations in physical characteristics. Suppl. to the Eugenics Soc. Bull. 2: 1-45.

ROGERS J.S. 1972: Measures of genetic similarity and genetic distance. Univ. Texas Publ. 7213: 145-153.

ROTHMAN E.D., NEEL J.V. & HOPPE F.M. 1981: Assigning a probability for paternity in apparent cases of mutation. Am. J. Hum. Genet. 33: 617-628.

RUCKNAGEL D.L. & WINTER W.P. 1974: Duplication of structural genes for hemoglobin α and β chains in man. Ann. N.Y. Acad. Sci. 241: 80-92.

SALMON D.B. & BROCTEUR J. 1976: Probability of paternity. Am. J. Hum. Genet. 28: 622-625.

SALMON D.B. & BROCTEUR J. 1978: Probability of paternity exclusion when relatives are involved. Am. J. Hum. Genet. 30: 65-75.

SALMON D., SEGER J. & SALMON C. 1980: Expected and observed proportion of subjects excluded from paternity by blood phenotypes of a child and its mother in a sample of 171 families. Am. J. Hum. Genet. 32: 432-444.

SAMUEL C.K. 1949: The Indian house sparrow, Passer domesticus L. var. indicus Jardine and Selby, as a serious orchard and wheat pest in Baluchistan. Indian J. Ent. 11: 219-220.

SAPPINGTON J.N. 1975: Cooperative Breeding in the House Sparrow, Passer domesticus. Ph.D. Thesis, Mississippi State University.

SASSAMAN C. 1978: Mating systems in porcellionid isopods. Multiple paternity and sperm mixing in Porcellio scaber Latr. Heredity 41: 385-397.

SCHACHT L.E. & GERSHOWITZ H. 1961: Frequency of extramarital children as determined by blood groups. In: BROCKINGTON F., FRANCOIS J., GEDDA L., et al. (Eds.) Second Intl. Conf. Hum. Genet. pp.894-897. Excerpta Medica, Amsterdam.

SCHARFETTER C. 1978: Alleged vs. biologically possible paternity. Behav. Genet. **8**: 383-384.

SCHENKER T. 1979: Four protein polymorphisms in pigeon brain extracts detected by two-dimensional gel electrophoresis. J. Neurochem. **33**: 1105-1114.

SCHIFFERLI L. 1976: Factors Affecting Weight and Conditions in the House Sparrow, Particularly when Breeding. D.Phil. Thesis, Oxford University.

SCHIFFERLI L. 1978: Experimental modifications of brood size among house sparrows Passer domesticus. Ibis **120**: 365-369.

SCHMIDT E. 1972: [On the bird food of the barn owl, Tyto alba and the long eared owl, Asio otus in Hungary.] Ornis Fenn. **49**: 98-102.

SCHMUTZ S.M. & SCHMUTZ J.F. 1981: Inheritance of color phases in ferruginous hawks. Condor **83**: 187-189.

SCHWARTZ O.A. & ARMITAGE K.B. 1980: Genetic variation in social mammals: the marmot model. Science **207**: 665-667.

SCHWARTZ O.A. & ARMITAGE K.B. 1983: Problems in the use of genetic similarity to show relatedness. Evolution **37**: 417-420.

SEEL D.C. 1960: The behaviour of a pair of house sparrows while rearing young. British Birds **53**: 303-310.

SEEL D.C. 1966: Further observations on a pair of house sparrows rearing young. Bird Study **13**: 207-209.

SEEL D.C. 1968a: Clutch-size, incubation and hatching success in the house sparrow and tree sparrow Passer spp. at Oxford. Ibis **110**: 270-282.

SEEL D.C. 1968b: Breeding seasons of the house sparrow and tree sparrow Passer spp. at Oxford. Ibis **110**: 129-144.

SEEL D.C. 1969: Food, feeding rates and body temperature in the nestling house sparrow, Passer domesticus, at Oxford. Ibis **111**: 36-47.

SEEL D.C. 1970: Nestling survival and nestling weight in the house sparrow and tree sparrow Passer spp. at Oxford. Ibis **112**: 1-14.

SELANDER R.K. 1972: Sexual selection and dimorphism in birds. In: CAMPBELL B.G. (Ed.) Sexual Selection and the Descent of Man, 1871-1971 pp.180-230. Aldine, Chicago.

SELANDER R.K. & JOHNSTON R.F. 1967: Evolution in the House Sparrow I. Intrapopulation variation in North America. Condor **69**: 217-258.

SELVIN S. 1980: Probability of nonpaternity determined by multiple allele codominant systems. Am. J. Hum. Genet. **32**: 276-278.

SHAW C.R. & PRASAD R. 1970: Starch gel electrophoresis of enzymes - a compilation of recipes. Biochem. Genet. **4**: 297-320.

SHIELDS W.M. 1982: Philopatry, Inbreeding and the Evolution of Sex. State University of New York Press.

SHIELDS W.M. 1983: Optimal inbreeding and the evolution of philopatry. In: SWINGLAND I.R. & GREENWOOD P.J. (Eds.) The Ecology of Animal Movement pp.132-159. Oxford University Press.

SIBLEY C.G. 1970: A comparative study of the egg-white proteins of passerine birds. Bulletin of the Peabody Museum of Natural History, Yale University, No.32.

SIBLEY C.G, CORBIN K.W., AHLQUIST J.E. & FERGUSON A. 1974: Birds. In: WRIGHT C.A. (Ed.) Biochemical and Immunological Taxonomy of Animals pp.89-176. Academic Press, London.

SIMONS R.C. & CRAWFORD M.H. 1969: Determination of paternity in group-born pigtailed monkeys, Macaca nemestrina. In: CARPENTER C.R. (Ed.) Proc. 2nd Intl. Congress of Primatology, Atlanta, **1**: 254 Basel, Karger.

SIMWAT G.S. 1977: Studies on the nidification of House Sparrow, Passer domesticus (L.) in Ludhiana (Punjab). J. Res. Punjab Agric. Univ. **14**: 193-198.

SINGLETON G.R. & HAY D.A. 1983: The effect of social organization on reproductive success and gene flow in colonies of wild house mice, Mus musculus. Behav. Ecol. Sociobiol. **12**: 49-56.

SMITH C.A.B. 1970: A note on testing the Hardy-Weinberg law. Ann. Hum. Genet. **33**: 377-383.

SMITH D.G. 1980: Paternity exclusion in six captive groups of rhesus monkeys, Macaca mulatta. Am. J. Phys. Anthropol. **53**: 243-249.

SMITH J.N.M. & DHONDT A.A. 1980: Experimental confirmation of heritable morphological variation in a natural population of song sparrows. Evolution **34**: 1155-1158.

SMITH J.N.M. & ZACH R. 1979: Song sparrow variation and heritability. Evolution **33**: 460-469.

SMITH N.J.H. 1973: House sparrows, Passer domesticus, in the Amazon. Condor **75**: 242-243.

SMITH N.J.H. 1980: Further advances of House Sparrows into the Brazilian Amazon. Condor **82**: 109-111.

SODHI S.S. & SINGH I.P. 1970: Probable foci of salmonella infection on a cattle farm. Indian J. Anim. Sci. **40**: 617-621.

SOKAL R.R. & ROHLF F.J. 1981: Biometry. W.H.Freeman, San Francisco.

SPENCER R. (Ed.) 1976: The Ringer's Manual. British Trust for Ornithology, Tring.

STACEY P.B. 1982: Female promiscuity and male reproductive success in social birds and mammals. Amer. Natur. **120**: 51-64.

STRATIL A. 1967: The effect of iron addition to avian egg white on the behaviour of conalbumin fractions in starch gel electrophoresis. Comp. Biochem. Physiol. **22**: 227-233.

SUMMERS-SMITH D. 1954: Colonial behaviour in the house sparrow. British Birds **47**: 249-265.

SUMMERS-SMITH D. 1958: Nest-site selection, pair formation and territory in the house sparrow Passer domesticus. Ibis **100**: 190-203.

SUMMERS-SMITH D. 1963: The House Sparrow. Collins, London.

SUMMERS-SMITH D. 1984: The systematic position of the sparrows (Passer). In: A New Dictionary of Birds. 3rd. Edition (in press).

SVENSSON L. 1975: Identification Guide to European Passerines. Naturhistoriska Riksmuseet, Stockholm.

TAMARIN R.H. & KREBS C.J. 1969: Microtus population biology II. Genetic changes at the transferrin locus in fluctuating populations of two vole species. Evolution **23**: 183-211.

TEGETMEIER W.B. 1899: The House Sparrow, the Avian Rat: in Relation to Agriculture and Gardening, with Practical Suggestions for Lessening Its Numbers. Vinton, London.

THORNEYCROFT H.B. 1976: A cytogenetic study of the White-throated Sparrow, Zonotrichia albicollis (Gmélin). Evolution **29**: 611-621.

TRIVERS R.L. 1972: Parental investment and sexual selection. In: CAMPBELL B. (Ed.) Sexual Selection and The Descent of Man 1871-1971 pp.136-179. Williams & Wilkins, Baltimore.

TURNER M.E. & ANDERSON W.W. 1983: Multiple mating and female fitness in Drosophila pseudoobøscura. Evolution **37**: 714-723.

VALENTIN J. 1980: Exclusions and attributions of paternity: practical experiences of forensic genetics and statistics. Am. J. Hum. Genet. **32**: 420-431.

VAN DELDEN W. 1982: The alcohol dehydrogenase polymorphism in Drosophila melanogaster: selection at an enzyme locus. Evol. Biol. **15**: 187-222.

VAN NOORDWIJK A.J. & SCHARLOO W. 1981: Inbreeding in an island population of the great tit. Evolution **35**: 647-688.

VAN NOORDWIJK A.J., VAN BALEN J.H. & SCHARLOO W. 1980: Heritability of ecologically important traits in the great tit, Parus major. Ardea **68**: 193-203.

VAN NOORDWIJK A.J., VAN BALEN J.H. & SCHARLOO W. 1981: Genetic variation in the timing of reproduction in the great tit. Oecologia **49**: 158-166.

VAURIE C. 1956: Systematic notes on palearctic birds No.24 Ploceidae: the genera Passer, Petronia and Montifringilla. Am. Mus. Nov. **1814.**

VOELKER R.A., LANGLEY C.H., LEIGH BROWN A.J. et al. 1980: Enzyme null alleles in natural populations of Drosophila melanogaster: Frequencies in a North Carolina population. Proc. Natl. Acad. Sci. USA **77:** 1091-1095.

VOHS P.A. & CARR L.R. 1969: Genetic and population studies of transferrin polymorphism in ring-necked pheasants. Condor **71:** 413-417.

WALD A. 1947: Sequential Analysis. Wiley, New York.

WALLACE B. 1958: The average effect of radiation-induced mutations on viability in Drosophila melanogaster. Evolution **12:** 532-552.

WALTON K.E., STYER D. & GRUENSTEIN E.I. 1979: Genetic polymorphism in normal human fibroblasts as analyzed by 2-dimensional polyacrylamide gel electrophoresis. J. Biol. Chem. **54:** 7951-7960.

WARD R.H. & SING C.F. 1970: A consideration of the chi-square test to detect inbreeding effects in natural populations. Amer. Natur. **104:** 355-365.

WASHINGTON D. 1973: Breeding the house sparrow. Avicultural Magazine **79:** 109-115.

WEAVER R.L. 1942: Growth and development of English sparrows. Wilson Bull. **54:** 183-191.

WEAVER R.L. 1943: Reproduction in English sparrows. Auk **60:** 62-74.

WEIR B.S. 1979: Inferences about linkage disequilibrium. Biometrics **35:** 235-254.

WEIR B.S. & COCKERHAM C.C. 1979: Estimation of linkage disequilibrium in randomly mating populations. Heredity **42:** 105-111.

WEIR B.S. & COCKERHAM C.C. 1984: Estimating F-statistics for the analysis of population structure. Evolution (In press).

WERSCHKUL D.F. 1982: Nesting ecology of the little blue heron: promiscuous behaviour. Condor **84**: 381-384.

WHITEHOUSE D.B. 1979: The Ecological Genetics of the Manx Shearwater (Aves) - A Preliminary Investigation. Ph.D. Thesis, Nottingham University.

WICKLER W. & SEIBT U. 1983: Monogamy: an ambiguous concept. In: BATESON P. (Ed.) Mate Choice pp.33-50. Cambridge University Press.

WILL R.L. 1970: Fecundity, density and movements of a house sparrow population in Southern Illinois. Ph.D. Thesis, University of Illinois.

WILLS C. 1978: Rank-order selection is capable of maintaining all genetic polymorphisms. Genetics **89**: 403-417.

WILLS C. 1980: Genetic Variability. Clarendon Press, Oxford.

WILSON J. 1981: Estimating the degree of polyandry in natural populations. Evolution **35**: 664-673.

WITSCHI E. 1935: Seasonal sex characters in birds and their hormonal control. Wilson bull. **47**: 177-188.

WITTENBERGER J.F. 1979: The Evolution of Mating Systems in Birds and Mammals. In: MARLER P. & VANDENBERGH J.G. (Eds.) Handbook of Behavioural Neurobiology Vol.3 pp.271-349. Plenum Press, New York.

WITTENBERGER J.F. 1981: Animal Social Behaviour. Duxbury Press, Boston.

WRIGHT S. 1922: Coefficients of inbreeding and relationship. Amer. Natur. **56**: 330-338.

WRIGHT S. 1951: The genetical structure of populations. Ann. Eugen. (Lond.) **15**: 323-354.

WRIGHT S. 1965: The interpretation of population structure by F-statistics with special regard to system of mating. Evolution **19**: 395-420.

WRIGHT S. 1979: Evolution and the Genetics of Populations. IV. Variability Within and Among Natural Populations. University of Chicago Press.

WYLLIE I. 1981: The Cuckoo. Batsford, London.

YALDEN D.W. 1980: Notes on the diet of urban kestrels. Bird Study **27**: 235-238.

YASUDA N. 1969: Estimation of the inbreeding coefficient from mating type frequency and gene frequency. In: MORTON N.E. (Ed.) Computer Applications in Genetics. University of Hawaii Press, Honolulu.

YOM-TOV Y. 1980a: Intraspecific nest parasitism among Dead Sea sparrows, Passer moabiticus. Ibis **122**: 234-237.

YOM-TOV Y. 1980b: Intraspecific nest parasitism in birds. Biol. Rev. **55**: 93-108.

YOM-TOV, Y., DUNNET G.M. & ANDERSON A. 1974: Intraspecific nest parasitism in the starling, Sturnus vulgaris. Ibis **116**: 87-90.

ZENONE P., SIMS M.E. & ERICKSON C.J. 1979: Male ring dove behaviour and the defense of genetic paternity. Amer. Natur. **114**: 615-626.

ZERA A.J., KOEHN R.K. & HALL J.G. 1983: Allozymes and biochemical adaptation. In: KERKUT G.A. & GILBERT L.I. (Eds.) Comprehensive Insect Physiology, Biochemistry and Pharmacology Pergamon Press, New York.

ZINK R.M. & BARROWCLOUGH G.F. 1984: Allozymes and song dialects: a reassessment. Evolution **38**: 444-448.

ZWICKEL F.C., REDFIELD J.A. & KRISTENSEN J. 1977: Demography, behaviour and genetics of a colonizing population of blue grouse. Can. J. Zool. **55**: 1948-1957.